

Scoring System for the Japanese Version of the ICECAP-A

Mitsuhiro Sado¹⁻³, Kengo Nagashima⁴, Akihiro Koreki^{2,3,5}

¹Keio University Health Center, Shinjuku-ku, Tokyo, Japan; ²Department of Neuropsychiatry, Keio University School of Medicine, Tokyo, Japan; ³Keio University Mindfulness & Stress Research Center, Tokyo, Japan; ⁴Biostatistics Unit, Clinical and Translational Research Center, Keio University Hospital, Tokyo, Japan; ⁵Department of Psychiatry, National Hospital Organization Shimofusa Psychiatric Medical Center, Chiba, Japan

Correspondence: Mitsuhiro Sado, Keio University Health Center, 35 Shinanomachi, Shinjuku-ku, Tokyo, 160-8582, Japan, Tel +81-3-5363-3235, Fax +81-3-5315-4349, Email mitsusado@keio.jp

Purpose: This study aimed to develop a value set for the Japanese version of the Investigating Choice Experiments CAPability Measure for Adults (ICECAP-A).

Patients and Methods: A total of 400 participants were recruited. Survey 1, conducted utilizing a self-report format, collected demographic data and responses to well-being scales, including the ICECAP-A, the Investigating Choice Experiments CAPability Measure for Older People (ICECAP-O), and the EuroQol 5 Dimensions (EQ-5D), among others. Participants who completed Survey 1 were invited to participate in Survey 2. In Survey 2, an interview-based assessment, participants engaged in best–worst scaling, where they identified the most and least favorable situations in each of the 16 hypothetical scenarios. A mixed-mixed multinomial logit (MM-MNL) model was utilized to estimate preference parameters, accounting for individual heterogeneity and latent classes.

Results: The estimated parameters and tariffs indicated that all five attributes contributed to an individual's capability for well-being, consistent with prior studies. The attributes were ranked in order of importance as follows: “enjoyment” (24.8% of the space), “attachment” (22.1%), “stability” (19.5%), “autonomy” (17.9%), and “achievement” (15.6%). For “enjoyment” and “attachment”, the differences between levels 1 and 2 and between levels 2 and 3 were equally large. In contrast, for all other attributes, the greatest differences were observed between levels 1 and 2. Across all attributes, the differences between levels 3 and 4 were comparatively smaller.

Conclusion: We developed a Japanese version of the ICECAP-A and demonstrated the feasibility of utilizing best–worst scaling in a non-European context. This approach allows for a more precise evaluation of quality of life (QoL) among individuals with chronic conditions, caregivers, and the general population. Future research should assess the scale's sensitivity in capturing changes in capability over time.

Keywords: quality of life, capability approach, investigating choice experiments capability measure for adults, ICECAP-A, cost-effective analysis, preference estimation, mixed-mixed multinomial logit model, MM-MNL model

Introduction

In many countries, rapidly aging populations and economic stagnation have placed significant constraints on financial resources allocated to healthcare. Consequently, economic evaluation in healthcare has emerged as an increasingly important international endeavor, playing a crucial role in optimizing the allocation of scarce resources.¹ Cost-utility analysis is commonly utilized to assess the efficiency of new healthcare technologies and interventions. This analysis typically relies on a narrow definition of health outcomes in terms of Quality-Adjusted Life Years (QALYs).² QALYs are estimated utilizing health-related quality of life (QoL) measures, with the Euro Quality of Life Scale 5 dimensions (EQ-5D)³ serving as the gold standard for effectiveness evaluation. EQ-5D is a widely utilized measure that evaluates mobility, self-care, limitations in usual activities, pain and discomfort, and anxiety and depression. Such measures are particularly well-suited for evaluating health interventions in acute care settings. However, there is growing recognition that measuring QoL solely from a health perspective in economic evaluations is inadequate.⁴⁻⁶ Some researchers advocate for moving beyond the QALY

framework,^{7–10} as traditional health-related QoL measures fail to capture broader aspects of well-being. This limitation is particularly evident when evaluating the QoL of individuals with chronic functional disabilities, such as dementia, who may still maintain social connections and a meaningful life, or when evaluating the QoL of their caregivers.¹¹ Additionally, understanding the QoL of the general population, the majority of whom belong to the non-clinical population, is equally crucial. To address these gaps, several instruments based on the capability approach have been developed to evaluate well-being.^{12–22} Among these, the Investigating Choice Experiments CAPability index (ICECAP)¹⁵ is the most widely utilized in economic evaluations.² ICECAP is grounded in Sen's capability approach,²³ which offers an alternative to traditional utilitarian welfare economics by focusing on what individuals are able to be and do in their lives.² As a result, ICECAP has the potential to provide an alternative measure—Capability-Adjusted Life Years (CALYs)²⁴—which offers a broader evaluative framework than QALYs for evaluating the impact of healthcare interventions. Unlike many psychological scales that calculate total scores by summing individual item responses, ICECAP-A is a preference-based measure that derives scores from population-based preference tariffs. ICECAP comprises several versions, including ICECAP for adults (ICECAP-A),¹⁵ ICECAP for older people (ICECAP-O),^{16,25} ICECAP for Supportive Care (ICECAP-SCM),²¹ and ICECAP for Close Person Measure (ICECAP-CPM).²⁰ ICECAP for children and young people (ICECAP-CYP) is currently being developed.²⁶ ICECAP-A, originally developed in the UK, has been validated for both the general population and clinical groups^{27–29} and has been translated into multiple languages, including German,³⁰ French,³¹ and Chinese.³² The measure generates a capability value ranging from 0 (worst state) to 1 (best state) based on respondent preferences. While value sets are available in the UK and other countries,^{11,33,34} Japan currently lacks a corresponding value set despite having a translated version of ICECAP-A. Developing a value set for the Japanese ICECAP-A would expand the evaluative framework for healthcare interventions in Japan, enabling more comprehensive assessments of well-being. Therefore, this study aims to develop a value set for the Japanese version of ICECAP-A.

Materials and Methods

Overview

To develop the scoring system for the Japanese version of the ICECAP-A, we conducted a two-phase study: a web-based questionnaire survey (Survey 1) and a web-based interview (Survey 2).

Target Population

Participants were selected to approximate the demographic distribution of the Japanese population in terms of age, sex, education level, and region of residence. Eligible individuals had to meet all inclusion criteria and none of the exclusion criteria.

Participation Criteria

Participants were required to meet the following inclusion criteria: 1) Japanese citizens from the general population, 2) aged 18 years or older, 3) able to understand the purpose of the survey and provide informed consent to participate in the survey, and 4) capable of completing the survey online, including participating in an online video conference interview.

Individuals who met the inclusion and exclusion criteria were excluded from the study: those aged 65 years or older who scored 20 or more points on the Dementia Self-Screening Checklist.³⁵

Sample Size

The required sample size for this study was 400. Although estimating an appropriate sample size requires data on the variability of preferences for each attribute or level, such data are unavailable in Japan. Therefore, the sample size was determined based on a prior study conducted in the UK, which required 400 complete responses.¹¹

Data Collected Through the Survey

The best–and worst scaling was obtained in Survey 2, while all other data was collected in Survey 1. Demographic data included age, sex, education level, region of residence, employment status, marital status, number of household members, household income, and past and present medical history.

The well-being-related scales utilized in this study were:

- Investigating Choice Experiments CAPability Measure for Adults (ICECAP-A)
- Investigating Choice Experiments CAPability Measure for Older People (ICECAP-O)
- EuroQol 5 Dimensions 5-Level (EQ-5D-5L)
- Flourishing Scale (FS)
- Satisfaction with Life Scale (SWLS)
- Scale of Positive and Negative Experiences (SPANE)
- Perceived Stress Scale (PSS)
- Perceived Health Status (PHS)

In best-worst scaling, participants identified the best and worst situations for each of the 16 hypothetical scenarios. Each scenario consisted of five attributes from the ICECAP-A—stability, attachment, autonomy, achievement, and enjoyment—each presented at one of four possible levels. The five levels within a scenario were selected from these attributes. An example scenario is exhibited in [Figure 1](#).

Survey Procedure

Surveys 1 and 2 were conducted by the CMIC Healthcare Institute Corporation (CHI Corporation), which has a survey panel of over eight million participants across various age groups and genders. Individuals meeting the inclusion criteria were randomly selected from this pool and invited to participate. Those willing to participate were directed to a screening page, where they reviewed an informed consent document prior to providing their consent. Participants who agreed to participate were then redirected to the survey platform through a website, where they completed the self-reported questionnaires for Survey 1. Participants who completed Survey 1 were subsequently invited to participate in Survey 2, which was scheduled later. During Survey 2, an interviewer met with each participant via an online videoconferencing system to provide instructions on completing the best–worst scaling task and to assist them as needed. The collected data,

Please look at this card and indicate which you think is the BEST thing in this situation.

Now we ask you to indicate which you think is the WORST thing in this situation.

Worst	Situation	Best
	I am able to feel settled and secure in a few areas of my life.	
	I cannot have any of love, friendship and support.	
	I am unable to be at all independent.	
	I can achieve and progress in a few aspects of my life.	
	I can have quite a lot of enjoyment and pleasure.	

Figure 1 A set of scenario examples. Each scenario consisted of five attributes from the ICECAP-A—stability, attachment, autonomy, achievement, and enjoyment—each presented at one of four possible levels. The five levels within a scenario were selected from these attributes.

excluding any personally identifiable information, were then sent to the Keio University Mindfulness and Stress Research Center for analysis.

Outcomes

The primary outcome of this study was to develop a tariff for ICECAP-A derived from the best–worst scaling results. The secondary outcomes included the mean and standard deviation of the collected variables, such as demographic data and well-being-related scales.

Instruments

The questionnaires utilized in Survey 1 are described below. All scales included in the study have been validated for utilization in Japan. Scores for ICECAP-A, ICECAP-O, and EQ-5D-5L, were calculated utilizing tariffs derived the preferences of the general population. For the remaining instruments, total scores were obtained by summing individual item scores.

Investigating Choice Experiments CAPability Measure for Adults

The ICECAP-A was devised to assess well-being capabilities in adults, an aspect not comprehensively captured by current health-related QoL measures. It comprises five attributes, each with four levels, producing a single index value that reflects overall well-being capability. Scores are derived utilizing tariffs informed by the preferences of the general population, ranging from 0 to 1. A higher score signifies a more favorable well-being status. In this study, the UK tariff was utilized.¹⁵

Investigating Choice Experiments CAPability Measure for Older People

The ICECAP-O is designed to evaluate well-being capability among older individuals. It encompasses five attributes: love and friendship, thinking about the future, doing things that make one feel valued, enjoyment and pleasure, as well as independence. Each attribute has four intensity levels, and capability scores are generated utilizing UK tariffs based on general population preferences. Scores range from 0 to 1, with higher values indicating better well-being.¹⁶

EuroQol 5 Dimensions 5-Level

The EQ-5D-5L is a standardized, preference-based measure to evaluate health-related QoL across diverse health conditions and treatments. It provides both a descriptive profile and a single index value representing health status. Scores range from 0 (death) to 1 (perfect health), with utility values estimated utilizing the Japanese version of the tariff.^{3,36,37}

Satisfaction with Life Scale

The SWLS is a five-item self-report questionnaire that evaluates the cognitive aspects of subjective well-being. Each item is rated on a scale from 1 (strongly disagree) to 7 (strongly agree), yielding total scores ranging from 5 to 35. Higher scores indicate greater life satisfaction.^{38,39}

Flourishing Scale

The Flourishing Scale comprises eight items assessing key aspects of human functioning, including positive relationships, competence, significance, and a sense of purpose in life. Participants rated each item on a seven-point scale, ranging from 1 (strongly disagree) to 7 (strongly agree). Total scores range from 8 (indicating strong disagreement across all items) to 56 (indicating strong agreement with all items). Higher scores reflect a more positive perception of overall well-being and personal functioning.^{40,41}

Scale of Positive and Negative Experience

The SPANE is a 12-item scale designed to assess both positive (six items) and negative experiences (six items). Due to its broad scope, the scale can gauge both pleasant and unpleasant feelings typically targeted by most scales, while also encompassing other conditions such as interest, flow, positive engagement, and physical pleasure. The SPANE-Positive (SPANE-P) and SPANE-Negative (SPANE-N) scales each range from 6 to 30, with higher scores indicating stronger

positive or negative affective states. The SPANE-Balance (SPANE-B) score, obtained by subtracting the negative score from the positive score, ranges from -24 to 24, with higher values reflecting a more positive emotional balance.^{40,41}

Perceived Stress Scale

The PSS was developed to assess the extent to which individuals perceive situations in their lives as stressful. It exists in two versions: the 14-item version (PSS-14) and the 10-item version (PSS-10), which closely resembles the PSS-14, which omits four items from the original scale. In this study, the PSS-10 was utilized to evaluate participants' perceived stress levels over the preceding month. Each item was rated on a five-point Likert scale ranging from 0 ("very often") to 4 ("never") to denote the frequency of positive experiences or responses. Total scores range from 0 to 40, with higher scores indicating elevated stress levels.^{42,43}

Perceived Health Status

The PHS scale was utilized to evaluate subjective health status. Employing a four-point Likert scale, ranging from 1 ("not healthy") to 4 ("very healthy"), it serves as a component of the national survey administered by the Ministry of Health, Labour and Welfare of Japan.⁴⁴

Adverse Events and Compensation

As this study was survey-based, no adverse events were anticipated. Participants who completed both the questionnaire and web-based interview were compensated with an Amazon gift card valued at JPY 3000.

Statistical Analysis for the Primary Outcome

All participants that provided complete answers were included in the analysis.

The primary outcome of this study was the development of a tariff for the ICECAP-A. Employing an orthogonal main-effects plan (OMEP) design, 16 scenarios were formulated, drawing from five attributes, each featuring four levels ([Supplementary Table 1](#)).¹¹ Within the OMEP framework, an equal allocation of each attribute level was ensured to facilitate the efficient estimation of the main effect. The participants were randomly assigned in a 1:1 ratio to either the original 16-scenario design described above or the folded-mirror version.

Best-minus-worst scores (BWS): The BWS were calculated as described by Flynn et al,¹¹ and scores were defined as the number of times each attribute level was selected as best minus, the number of times it was selected as the worst. For the five attributes, the sum of squares (SS) of the normalized (ie, divided by four) BWS and the empirical scale parameter (ESP), that is, the sum of the SSs, were also calculated. Summary statistics of the SSs and ESP were calculated, and a histogram of the ESP was obtained.

Methods for estimating tariffs: A mixed-mixed multinomial logit (MM-MNL) model was then utilized to estimate the preference parameters for individual-level heterogeneity and latent classes. In the MM-MNL model, the choice probabilities of the responder i ($i=1, \dots, n$) for the attribute j ($j=1, \dots, 5$) as the best or worst k ($k=1, 2$) in scenario l ($l=1, \dots, 16$) are given by:

$$P_i(\theta) = \sum_{q=1}^Q w_{iq} \int \left\{ \prod_{j=1}^J \prod_{k=1}^K \prod_{l=1}^L \left[\frac{\exp(\mathbf{x}_{ijkl}^T \boldsymbol{\beta}_i)}{\sum_m \exp(\mathbf{x}_{imkl}^T \boldsymbol{\beta}_i)} \right]^{y_{ijkl}} \right\} f_q(\boldsymbol{\beta}_i) d\boldsymbol{\beta}_i,$$

where $\boldsymbol{\theta} = (\boldsymbol{\beta}_1, \dots, \boldsymbol{\beta}_q, \boldsymbol{\Sigma}_1, \dots, \boldsymbol{\Sigma}_q, \gamma_1, \dots, \gamma_q)^T$, $\boldsymbol{\beta}_i$ follows a q -variate normal distribution $N(\boldsymbol{\beta}_q, \boldsymbol{\Sigma}_q)$ with probability w_{iq} for a latent class $q=1, \dots, Q$, $f_q(\boldsymbol{\beta}_i) = N(\boldsymbol{\beta}_q, \boldsymbol{\Sigma}_q)$, $\boldsymbol{\beta}_q = (\beta_{q,1}, \dots, \beta_{q,20})^T$, are coefficients that correspond to each level of each attribute (ie, $\beta_{q,1}$ is the coefficient of "stability" Level 4, ..., and $\beta_{q,20}$ is the coefficient of "enjoyment" Level 1 for latent class q), \mathbf{x}_{ijkl} is a design vector of regression coefficients that are effect coded, if selected $y_{ijkl} = 1$ and otherwise $y_{ijkl} = 0$, $w_{iq} = \exp(\mathbf{h}_i^T \boldsymbol{\gamma}_q) / \sum_{r=1}^Q \exp(\mathbf{h}_i^T \boldsymbol{\gamma}_r)$, $\gamma_1 = 0$, $\mathbf{h}_i = (1, SS_{1i}, SS_{2i}, SS_{3i}, SS_{4i}, SS_{5i})^T$ is a SSs design vector of mixture probabilities for responder i , $\boldsymbol{\gamma}_q$ is the parameter vector for mixture probabilities w_{qi} , and superscript "T" denotes matrix transposition.

From the estimates of the above model, the tariffs were calculated as follows. Based on the linear combination of the 20 parameters estimated by the weighted average, $\beta = \sum_{q=1}^Q w_q \beta_q$, a tariff was obtained with the state of “no capability” (1111) as 0 and the state of “full capability” (4444) as 1, where $w_q = \frac{\sum_i w_{iq}}{n}$. A scatter plot of utilities computed utilizing the UK and Japanese tariffs was drawn.

Sensitivity analyses were conducted to address uncertainties associated with the choice of statistical model and to confirm the robustness of the estimated results. Four different models were evaluated, each with different assumptions: a scale-adjusted multinomial logit model adjusted by ESP, a mixed-mixed logit model without class share terms, a latent class logit model, and a multinomial logit model. The results from these models were compared with those of the primary model described above.

Software

All statistical analyses were conducted utilizing R software (version 4.2.3; R Foundation for Statistical Computing, Vienna, Austria) with the support of the support.BWS2,⁴⁵ mlogit,⁴⁶ gmn1,⁴⁷ and gtsummary⁴⁸ packages.

Results

Characteristics of the Participants

A total of 400 participants were included in the analysis. The main characteristics of the participants are presented in Table 1, with additional details provided in [Supplementary Table 2](#). The mean age of the participants was 51 years, and 50.5% were female. Although comparisons with the national data^{49,50} indicate slight differences in certain attributes (eg, age category: 18–34; education: primary/junior high), these differences were considered minor, and the sample was deemed sufficiently representative.

The Results of the Wellbeing Questionnaires

The results of the well-being questionnaires are summarized in Table 2. The mean capability/ utility scores (standard deviation) estimated utilizing the ICECAP-A, ICECAP-O and EQ-5D-5L were 0.828 (0.170), 0.748 (0.178) and 0.891 (0.113), respectively. The results for subjective health indicated that the proportion of participants rating their health as “very healthy” or “rather healthy” (72.8%) was comparable to that reported in the national survey (73.7%).⁴⁴

Best–Worst Pairs

The results of the best-worst pairs from the best-worst scaling analysis are presented in Table 3. Among the attributes, “enjoyment” was most frequently selected as the “best choice” (835 frequencies), followed by “stability” (721 frequencies), “attachment” (679 frequencies), “achievement” (425 frequencies), and “autonomy” (321 frequencies). Regarding the “worst choice”, which represents the attribute participants most desired to avoid, “enjoyment” was again the most frequently selected (1133 frequencies), followed by “attachment” (933), “autonomy” (780), “stability” (652), and “achievement” (412).

Best-Minus-Worst Score (BWS)

The results for each SS of the normalized BWS and ESP, which represents the sum of the SSs, are provided in Table 4. The mean EPS was 4.49 (SD = 0.44), suggesting that participants understood the survey well and exhibited consistent preferences. The averages SS values for SS5 (1.34) and SS2 (1.05) were higher than those for other attributes, indicating that respondents considered “enjoyment” and “attachment” particularly important.

Regression Estimates and Tariffs

We evaluated both two-preference and three-preference class models as potential candidates. Model fit was assessed utilizing the Bayesian Information Criterion (BIC). The two-preference class model yielded a BIC value of 22,355.29, whereas the three-preference class model produced a BIC value of 22,414.47. Given that a lower BIC value indicates a better fit, the two-preference class model was deemed superior. Consequently, we selected the two-preference class

Table 1 Demographic Characteristics

Characteristic	Sample (n = 400)		National Data	
	n	%	%	Source
Age category				Statistics of bureau of Japan*1
18–34	92	23.0	19.2	
35–49	95	23.8	22.9	
50–64	84	21.0	23.5	
65+	129	32.3	34.5	
Gender				Statistics of bureau of Japan*1
Male	198	49.5	48.6	
Female	202	50.5	51.4	
Education				Statistics of bureau of Japan*2
Primary/junior high	44	11.0	14.1	
High school	179	44.8	44.2	
Junior college/technical college	75	18.8	16.2	
Undergraduate	94	23.5	25.6	
Postgraduate (masters and doctorate)	8	2.0	2.4	
Regions of residence				Statistics of bureau of Japan*1
Hokkaido/Tohoku	43	10.8	11.0	
Kanto/Koshinetsu	155	38.8	38.6	
Tokai/Hokuriku	61	15.3	14.0	
Kansai	66	16.5	16.3	
Chugoku/Shikoku	33	8.3	8.7	
Kyushu	42	10.5	11.4	
Employment				
Full time	120	30.0		
Part time	65	16.3		
Pensioner	45	11.3		
Students	28	7.0		
Unemployed	65	16.3		
Home maker	77	19.3		
Marital status				
Married	230	57.5		
Partnership	5	1.3		
Single	108	27.0		
Divorced	46	11.5		
Widow/ widower	11	2.8		
Families living together (adults)		0.0		
None	82	20.5		
1 adult	179	44.8		
2 adults	89	22.3		
3 or more adults	50	12.5		
Families living together (minors)				
None	327	81.8		
1 minor	41	10.3		
2 minors	23	5.8		
3 or more minors	5	1.3		
Unknown	4	1.0		
Annual net household income (million JPY)				
< 2 million JPY	60	15.0		
2 <= 4 million JPY	101	25.3		
4 <= 6 million JPY	90	22.5		

(Continued)

Table 1 (Continued).

Characteristic	Sample (n = 400)		National Data	
	n	%	%	Source
6 <= 8 million JPY	79	19.8		
8 <= 10 million JPY	33	8.3		
10 <= 12 million JPY	15	3.8		
12 <= <15 million JPY	13	3.3		
15 <= 20 million JPY	6	1.5		
20 million JPY+	3	0.8		

Notes: *1 Statistics of Bureau of Japan: Population Estimates: Statistics of Japan 2022. Tokyo, Statistics of Bureau of Japan; 2023. *2 Statistics of Bureau of Japan: Population Census 2020 Population Census Basic Complete Tabulation on Labour Force. Tokyo, Statistics of Bureau of Japan; 2022.

Table 2 Secondary Outcomes

Characteristic	N = 400 ¹
ICECAP-A	0.828 (0.170)
ICECAP-O	0.748 (0.178)
EQ-5D-5L ²	0.891 (0.113)
EQ-5D-5L: VAS	79.8 (18.3)
FS	38.4 (9.0)
SWLS	20.6 (7.3)
SPANE	5.5 (9.9)
PSS	23.5 (7.4)
Subjective health status	
Not healthy	24 (6.0%)
Not very healthy	85 (21.2%)
Rather healthy	232 (58.0%)
Very healthy	59 (14.8%)

Notes: ¹ Mean (SD); n (%). ² N = 396 (excluding six observations with missing data).

Abbreviations: ICECAP-A, Investigating Choice Experiments CAPability Measure for Adults; ICECAP-O, Investigating Choice Experiments CAPability Measure for Older People; EQ-5D-5L, EuroQol 5 Dimensions 5 levels; EQ-5D-5L, VAS, EuroQol 5 dimensions visual analog scale; FS, Flourishing Scale; SWLS, Satisfaction with Life Scale; SPANE, Scale of Positive and Negative Experience; PSS, Perceived Stress Scale; PHS, Perceived Health Status.

model as the final estimate, aligning with findings from prior studies.^{11,34} The tariffs estimated in the three-preference class model were similar.

The regression parameter estimates and tariffs are presented in Table 5. The estimated parameters suggest that all five attributes contribute to an individual's capability for well-being, consistent with prior research.¹¹ The attributes ranked by importance were as follows: “enjoyment” (24.8% of the space), “attachment” (22.1%), “stability” (19.5%), “autonomy” (17.9%), and “achievement” (15.6%). For the attributes “enjoyment” and “attachment”, the differences between levels 1 and 2 and levels 2 and 3 were equally large. For all attributes except “enjoyment” and “attachment”, the greatest differences were observed between levels 1 and 2, whereas the differences between levels 3 and 4 were much smaller.

The tariffs for each latent class and the estimated class membership parameters are provided in Table 5. The class shares were 77.2% for latent class 1 and 22.8% for latent class 2. In latent class 1, the order of importance was “enjoyment”

Table 3 Best-Worst Pair Frequencies (n = 399, 16 Scenarios), ICECAP-A

Best attribute level (rows)		Worst Attribute Level (columns)																				Total Best
		Stability				Attachment				Autonomy				Achievement				Enjoyment				
		1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4	
Stability	1					3	0	0	0	0	0	0	0	1	0	0	0	1	1	0	0	6
	2					22	9	2	0	42	6	5	11	9	0	4	9	43	20	3	2	187
	3					162	25	2	2	50	42	3	20	62	31	3	2	166	40	0	1	611
	4					113	30	18	0	77	115	37	6	42	11	18	2	135	108	2	7	721
Attachment	1	0	0	0	0					0	0	0	0	0	0	0	0	0	0	0	0	0
	2	13	1	0	1					10	1	0	2	6	1	3	5	29	0	3	1	76
	3	62	5	2	13					132	70	5	9	83	13	6	9	143	50	2	5	609
	4	101	46	3	4					94	48	18	8	37	52	8	20	174	61	3	2	679
Autonomy	1	0	0	0	0	0	0	0	0					0	0	0	0	0	0	0	0	0
	2	7	0	1	3	7	2	3	4					15	0	1	3	35	3	2	1	87
	3	74	15	3	6	99	23	4	7					25	15	7	8	56	30	3	5	380
	4	32	6	3	4	24	78	4	3					26	5	3	2	69	57	2	3	321
Achievement	1	1	0	0	0	0	0	0	0	2	0	0	0					0	0	0	0	3
	2	48	3	3	2	34	0	4	3	19	1	1	6					112	8	0	2	246
	3	27	12	2	3	91	44	7	0	47	15	0	3					93	72	0	1	417
	4	37	52	0	3	45	76	1	2	68	48	2	0					77	13	1	0	425
Enjoyment	1	1	0	0	0	0	1	0	0	0	0	0	0	1	0	0	0					3
	2	6	1	5	2	11	3	2	3	9	2	2	4	4	2	3	5					64
	3	139	34	4	5	152	52	3	0	124	100	8	15	43	10	5	20					714
	4	104	38	4	17	170	87	9	1	106	79	46	15	58	73	23	5					835
Total worst		652	213	30	63	933	430	59	25	780	527	127	99	412	213	84	90	1133	463	21	30	6384

Table 4 Summary Statistics of ESP and SS

Scores	N = 399
ESP	4.49 (0.44)
SS1 (stability)	0.85 (0.69)
SS2 (attachment)	1.05 (0.75)
SS3 (autonomy)	0.74 (0.62)
SS4 (achievement)	0.51 (0.58)
SS5 (enjoyment)	1.34 (0.73)

Note: * Values are shown as means (SD).

Abbreviations: SS, sum of squares of normalized (ie, divided by four) BVVS scores; ESP, empirical scale parameter.

(24.6%), “attachment” (22.1%), “stability” (19.8%), “autonomy” (17.7%), and “achievement” (15.8%). Latent class 2 had a slightly different ranking: “enjoyment” (26.9%), “attachment” (22.3%), “autonomy” (20.0%), “stability” (16.9%), and “achievement” (14.0%). The estimated class membership parameters indicated that respondents with lower SSs or lower EPS were more likely to belong to latent class 2. Given the consistency in the ordering of regression parameter estimates within latent class 2, this class may exhibit slightly inconsistent preferences. However, because latent class 2 represents a smaller proportion of the sample, the contribution of latent class 1 to the overall tariff is considered more significant.

Table 5 Estimated Preference Parameters, Implied Tariffs, and Class Membership Parameters by the Mixed-Mixed Logit Model

	Latent Class 1		Latent Class 2		Japanese Tariff			UK Tariff
	Estimate	SE	Estimate	SE	Latent Class 1	Latent Class 2	Overall	
Preference parameters and tariffs								
Stability (mean)	0.4424	0.1464	−0.1266	0.1044				
Attachment (mean)	0.0940	0.1552	−0.1624	0.0942				
Autonomy (mean)	0.3203	0.1409	0.0267	0.0986				
Enjoyment (mean)	0.3079	0.1422	−0.1966	0.0983				
Stability (4)	3.6226	0.1271	0.4965	0.1389	0.1987	0.1775	0.1966	0.2221
Stability (3)	2.5683	0.1074	1.1938	0.1072	0.1767	0.2162	0.1807	0.1915
Stability (2)	−0.2962	0.1114	0.8588	0.1330	0.1172	0.1976	0.1252	0.1013
Stability (1)	−5.8947	0.2361	−2.5490	0.1471	0.0007	0.0083	0.0015	−0.0008
Attachment (4)	4.1930	0.1340	1.1871	0.1353	0.2105	0.2158	0.2110	0.2276
Attachment (3)	3.5496	0.1232	1.4121	0.1172	0.1971	0.2283	0.2002	0.1890
Attachment (2)	−1.3109	0.1151	0.2293	0.1113	0.0961	0.1626	0.1027	0.0964
Attachment (1)	−6.4317	0.2482	−2.8285	0.1710	−0.0104	−0.0072	−0.0101	−0.0239
Autonomy (4)	3.2884	0.1236	0.9042	0.1826	0.1917	0.2001	0.1925	0.1881
Autonomy (3)	2.5215	0.1013	1.5054	0.1084	0.1758	0.2335	0.1815	0.1560
Autonomy (2)	−0.5861	0.1026	0.2800	0.1558	0.1111	0.1655	0.1165	0.0836
Autonomy (1)	−5.2238	0.2332	−2.6896	0.1599	0.0147	0.0005	0.0133	0.0063
Achievement (4)	2.6504	0.1240	0.3497	0.1399	0.1784	0.1693	0.1775	0.1811
Achievement (3)	2.2272	0.1027	0.6300	0.1123	0.1696	0.1849	0.1712	0.1588
Achievement (2)	0.0677	0.1014	1.1867	0.1162	0.1247	0.2158	0.1338	0.0909
Achievement (1)	−4.9454	0.2352	−2.1664	0.1636	0.0205	0.0296	0.0214	0.0210
Enjoyment (4)	4.6811	0.1657	1.5725	0.1221	0.2207	0.2372	0.2223	0.1811
Enjoyment (3)	3.9548	0.1641	1.5853	0.1081	0.2056	0.2380	0.2088	0.1540
Enjoyment (2)	−1.4822	0.1569	0.1057	0.1113	0.0925	0.1558	0.0988	0.0693
Enjoyment (1)	−7.1537	0.2669	−3.2635	0.1433	−0.0255	−0.0313	−0.0260	−0.0026
Class membership parameters								
Intercept	—	—	5.9966	0.2484				
SS1 (stability)	—	—	−1.9907	0.0679				
SS2 (attachment)	—	—	−1.8510	0.0660				
SS3 (autonomy)	—	—	−0.7387	0.0630				
SS4 (achievement)	—	—	−1.8630	0.0746				
SS5 (enjoyment)	—	—	−1.7458	0.0625				
Class share	0.7724		0.2276					

Discussion

Overall Results

We developed a Japanese version of the ICECAP-A tariff. As previously noted, while health-related QoL assessments may not comprehensively capture all pertinent aspects of QoL,^{4–6} the introduction of the ICECAP-A—a capability based scale rooted in the capability approach—is a significant advancement. This scale enabled a more precise evaluation of QoL beyond QALYs, particularly for individuals with chronic conditions, caregivers, and the general public. Although the methodology for applying capability scores within the capability approach remains an ongoing subject of debate, we anticipate that the increasing adoption of ICECAP-A in Japan will provide a richer framework for assessing patients' QoL.

An analysis of frequency distributions from the best-worst scaling revealed that “enjoyment” was the most highly valued attribute, followed by “stability” and “attachment”, whereas “achievement” ranked lowest, followed by “autonomy”. While the ranking slightly shifted when considering attributes with the highest value based on the best–worst

scoring (ie, enjoyment, attachment, and stability), the consistent prioritization of enjoyment suggests its central importance.

We identified a small degree of heterogeneity in preferences within our sample, as evidenced by the detection of two latent classes. In the second latent class, individuals place greater emphasis on enjoyment and attachment, while assigning comparatively lower importance to achievement and stability than the overall sample. Consequently, there was a disparity in tariff scores between the most valued attribute (enjoyment: 0.2372) and the least valued attribute (achievement: 0.1693) (Table 5). Conversely, individuals in the first latent class exhibit a more balanced emphasis on all attributes. Hence, the gap in tariff scores between the most valued (enjoyment: 0.2207) and least valued (achievement: 0.1784) groups was narrower. As exhibited in the Results section, the greatest increase in capability occurred when transitioning from level 1 to level 2 across all attributes. As attribute levels increased, the magnitude of capability improvement diminished, with the smallest grains observed when moving from level 3 to level 4. This pattern suggests that individuals place greater value on escaping from the worst state than on transitioning from a better to the best state.

Feasibility

We adopted the same methodology utilized in prior studies^{11,33,34} and demonstrated its applicability beyond the European context. Additionally, we found that the best–worst scaling exercise was feasible and could serve as a viable alternative to the discrete choice experiment, as interviewers reported no notable difficulties in assisting respondents with completing the exercise.

Comparison with Other Countries

A comparison of the results with those from European countries highlights two key differences. The first, concerns the ranking of the most valued attributes. In both the UK and Hungary, “attachment” was ranked as the most important attribute, followed by “stability” and “enjoyment”. In contrast, in Japan, “enjoyment” was the most valued attribute, followed by “stability” and “attachment”. Although this study does not directly investigate the reasons for this difference, Japan’s cultural background may have influenced these preferences. In ICECAP-A, “attachment” encompasses love, friendship, and support. In Japanese culture, where individuals are primarily perceived as members of a social group, indirect communication is often preferred—both when giving and refusing directives—to maintain group harmony.⁵¹ Given this emphasis on implicit understanding, affection, friendship, and mutual support may be perceived as inherently understood rather than consciously acknowledged, potentially contributing to their relatively lower explicit evaluation.

The second key difference relates to the degree of capability improvement across different levels. As exhibited in [Supplementary Table 3](#), although both Japan and the UK exhibited a general trend where capability improvement diminished as levels increased, the decline was more pronounced in Japan. For instance, in Japan, capability increases from levels 1 to 2 exceeded 0.1 across all attributes, whereas increases from levels 3 to 4 were less than 0.02 for all attributes. In contrast, in the UK, only “stability” and “attachment” exhibited capability increases of more than 0.1 when transitioning from level 1 to 2, while increases from level 3 to 4 exceeded 0.2 across all attributes. This implies that Japanese respondents emphasize avoiding the worst state rather than acquiring the best compared to the general public in the UK. This pattern is further reflected in the scatterplot of utilities computed utilizing the UK and Japanese tariffs, where all data points are positioned on or above the slope of one (Figure 2). These differences highlight the importance of developing country- and culture-specific tariffs to accurately reflect preference variations.

Strengths and Limitations

This study has several strengths. First, we employed a methodology—best–worst scaling—that helps reduce the cognitive burden on respondents. Compared to a discrete choice experiment requiring a more complex understanding of trade-offs, best–worst scaling allows respondents to instinctively indicate their preferences. We believe the obtained data are consistent and reliable for the following reasons. First, the proportions of responses in which higher levels (levels 3 and 4) were selected as the “best” and lower levels (levels 1 and 2) as the “worst” were considerably high (89.5% and 90.2%, respectively). Second, as observed in a prior study conducted in Hungary, neither completely random responses nor responses indicating gaming behavior were detected in the ESP distribution (Figure 3). Third, to address

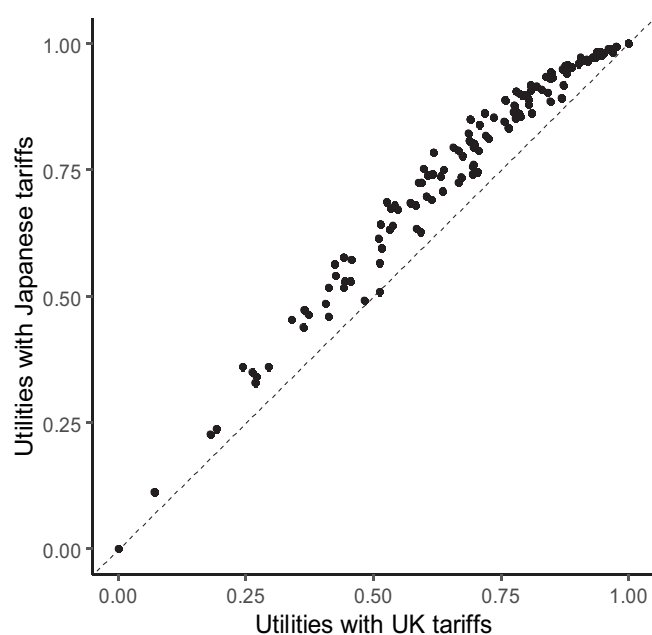


Figure 2 Scatterplot of utilities computed with UK and Japanese tariffs (N = 400). The dotted line represents a reference with a slope of one.

uncertainties related to model selection and to confirm the robustness of the study's estimates, we conducted sensitivity analyses. In addition to the final model, four alternative models were analyzed with different assumptions, all yielding similar results (see [Supplementary Table 4](#)).

Despite these strengths, this study has certain limitations. The first concerns the representativeness of the sample. Ideally, respondents should be randomly recruited from the general public. However, due to feasibility constraints, participants were recruited through a survey company with access to a large pooled sample of individuals in internet-

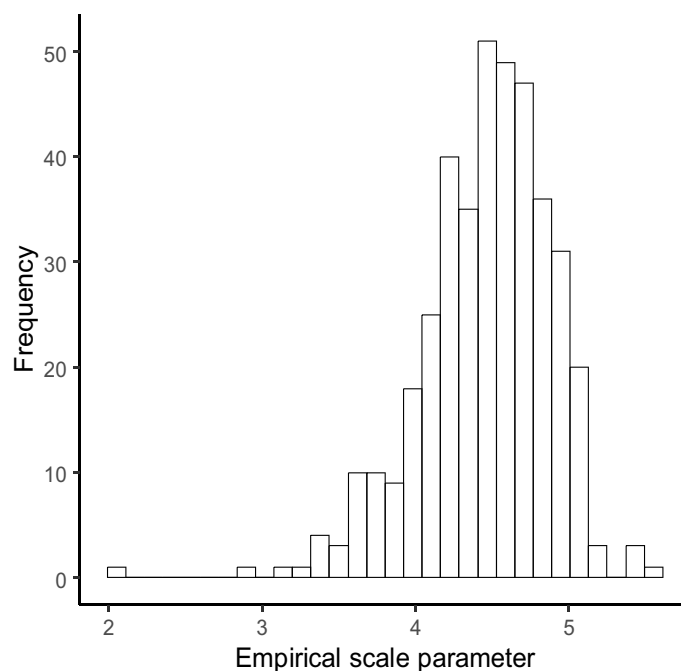


Figure 3 Empirical scale parameter distribution. Neither completely random responses nor responses indicating gaming behavior were detected.

enabled environments. Consequently, selection bias may be present. Furthermore, participants were voluntarily recruited from this pooled sample rather than through random sampling, introducing the possibility of sampling bias. That said, the demographic characteristics of this study's sample closely align with those of the general population, suggesting that any bias introduced is minimal and within an acceptable range. Future research should further evaluate the sensitivity and responsiveness of the ICECAP-A by examining score changes in relation to established clinical scales.

Conclusions

Following the methodology adopted in prior studies, we developed a Japanese version of the ICECAP-A and demonstrated the feasibility of best-worst scaling in Japan. The results indicate that “enjoyment” was the most valued attribute, followed by “stability” and “attachment”. Future research should assess whether the scale is sufficiently sensitive to capture changes respondents' capabilities over time.

Abbreviations

BIC, Bayesian Information Criterion; BWS, Best-minus-Worst Scores; CALYs, Capability-Adjusted Life Years; EQ-5D-5L – EuroQoL 5-Dimensions 5-Level scale; ESP, Empirical Scale Parameter; FS, Flourishing Scale; ICECAP-A – ICEpop CAPability Measure for Adults; ICECAP-SCM – ICEpop CAPability Measure for Social Care; ICECAP-CPM – ICEpop CAPability Measure for Palliative care; ICECAP-CYP – ICEpop CAPability Measure for Children and Young People; MM-MNL, Mixed-Mixed MultiNomial Logit; OMEP, Orthogonal Main-Effects Plan; PSS, Perceived Stress Scale; PHS, Perceived Health Status; QALYs, Quality-Adjusted Life Years; QoL, quality of life; SPANE – Scale of Positive and Negative Experience; SS, Sum of Squares; SWLS, Satisfaction with Life Scale; UK, United Kingdom.

Data Sharing Statement

Data are available in Keio University Mindfulness & Stress Research Center for researchers who meet the access criteria for sensitive data, as defined by the Ethics Committee of Keio University.

Ethical Approval and Informed Consent

This study was approved by the ethics committee of Keio University (ID: 20221077) and was conducted in accordance with the Declaration of Helsinki, as well as the ethical guidelines for life science and medical research involving human subjects published by the Japanese Ministry of Health, Labor and Welfare.

Acknowledgments

We sincerely thank Prof. Hideyuki Kobayashi, Prof. Reiko Goto, and Prof. Ryu Kambayashi for their valuable suggestions and insightful comments. We also appreciate Editage (www.editage.jp) for their assistance with English language editing.

Funding

This research was funded by Japan Research Institute Limited, as part of a joint research project with Keio University under the Ministry of Economy, Trade, and Industry's 2022 Social Implementation Project for Health Care Services. The Japan Research Institute, Inc., provided the necessary funding for this research and collaborates with Keio University as a research management organization. To ensure the integrity of the study, the Japan Research Institute, Inc., was not involved in any research activities, including data management, statistical analysis, or database access during the research period.

Disclosure

Prof. Dr. Mitsuhiro Sado has received honoraria from Shionogi & Co., Ltd., Meiji Seika Pharma, Hisamitsu Pharmaceutical, Mochida Pharmaceutical, and Terumo Corporation, as well as grants from Sony Corporate Services Corporation and Meiji Seika Pharma, outside of the scope of this work. The authors declared no other conflicts of interest in this work.

References

- Glick HA, Doshi JA, Sonnad SS, Polsky D. *Economic Evaluation in Clinical Trials*. Oxford: Oxford University Press; 2014.
- Helter TM, Coast J, Laszewska A, Stamm T, Simon J. Capability instruments in economic evaluations of health-related interventions: a comparative review of the literature. *Qual Life Res*. 2020;29:1433–1464. doi:10.1007/s11136-019-02393-5
- Brooks R. EuroQol: the current state of play. *Health Policy*. 1996;37:53–72. doi:10.1016/0168-8510(96)00822-6
- Oliver A, Healey A, Donaldson C. Choosing the method to match the perspective: economic assessment and its implications for health-services efficiency. *Lancet*. 2002;359:1771–1774. doi:10.1016/S0140-6736(02)08664-6
- Coast J. Is economic evaluation in touch with society's health values? *BMJ*. 2004;329:1233–1236. doi:10.1136/bmj.329.7476.1233
- Ryan M, Netten A, Skatun D, Smith P. Using discrete choice experiments to estimate a preference-based measure of outcome—an application to social care for older people. *J Health Econ*. 2006;25:927–944. doi:10.1016/j.jhealeco.2006.01.001
- Nord E. Beyond QALYs: multi-criteria based estimation of maximum willingness to pay for health technologies. *Eur J Health Econ*. 2018;19:267–275. doi:10.1007/s10198-017-0882-x
- Greco G, Lorgelly P, Yamabhai I. Outcomes in economic evaluations of public health interventions in low- and middle-income countries: health, capabilities and subjective wellbeing. *Health Econ*. 2016;25 Suppl 1:83–94. doi:10.1002/heec.3302
- Makai P, Brouwer WB, Koopmanschap MA, Stolk EA, Nieboer AP. Quality of life instruments for economic evaluations in health and social care for older people: a systematic review. *Soc Sci Med*. 2014;102:83–93. doi:10.1016/j.socscimed.2013.11.050
- Al-Janabi H, Flynn TN, Coast J. QALYs and carers. *Pharmacoeconomics*. 2011;29:1015–1023. doi:10.2165/11593940-000000000-00000
- Flynn TN, Huynh E, Peters TJ, et al. Scoring the Icecap-a capability instrument. Estimation of a UK general population tariff. *Health Econ*. 2015;24:258–269. doi:10.1002/heec.3014
- Sacchetto B, Aguiar R, Vargas-Moniz MJ, et al. The capabilities questionnaire for the community mental health context (CQ-CMH): a measure inspired by the capabilities approach and constructed through consumer-researcher collaboration. *Psychiatr Rehabil J*. 2016;39:55–61. doi:10.1037/prj0000153
- Simon J, Anand P, Gray A, Rugkasa J, Yeeles K, Burns T. Operationalising the capability approach for outcome measurement in mental health research. *Soc Sci Med*. 2013;98:187–196. doi:10.1016/j.socscimed.2013.09.019
- Netten A, Burge P, Malley J, et al. Outcomes of social care for adults: developing a preference-weighted measure. *Health Technol Assess*. 2012;16:1–166. doi:10.3310/hta16160
- Al-Janabi H, Flynn TN, Coast J. Development of a self-report measure of capability wellbeing for adults: the ICECAP-A. *Qual Life Res*. 2012;21:167–176. doi:10.1007/s11136-011-9927-2
- Grewal I, Lewis J, Flynn T, Brown J, Bond J, Coast J. Developing attributes for a generic quality of life measure for older people: preferences or capabilities? *Soc Sci Med*. 2006;62:1891–1901. doi:10.1016/j.socscimed.2005.08.023
- Turnpenny A, Caiels J, Whelton B, et al. Developing an easy read version of the adult social care outcomes toolkit (ASCOT). *J Appl Res Intellect Disabil*. 2018;31:e36–e48. doi:10.1111/jar.12294
- Rand S, Caiels J, Collins G, Forder J. Developing a proxy version of the adult social care outcome toolkit (ASCOT). *Health Qual Life Outcomes*. 2017;15:108. doi:10.1186/s12955-017-0682-0
- Botes R, Vermeulen KM, Gerber AM, Ranchor AV, Buskens E. Functioning and quality of life in Dutch oldest old with diverse levels of dependency. *Patient Prefer Adherence*. 2018;12:2187–2196. doi:10.2147/PPA.S175388
- Canaway A, Al-Janabi H, Kinghorn P, Bailey C, Coast J. Development of a measure (ICECAP-close person measure) through qualitative methods to capture the benefits of end-of-life care to those close to the dying for use in economic evaluation. *Palliat Med*. 2017;31:53–62. doi:10.1177/0269216316650616
- Sutton EJ, Coast J. Development of a supportive care measure for economic evaluation of end-of-life care using qualitative methods. *Palliat Med*. 2014;28:151–157. doi:10.1177/0269216313489368
- Lorgelly PK, Lorimer K, Fenwick EA, Briggs AH, Anand P. Operationalising the capability approach as an outcome measure in public health: the development of the OCAP-18. *Soc Sci Med*. 2015;142:68–81. doi:10.1016/j.socscimed.2015.08.002
- Sen A. *Inequality Reexamined*. London: Penguin; 1992.
- Mansdotter A, Ekman B, Feldman I, Hagberg L, Hurtig AK, Lindholm L. We propose a novel measure for social welfare and public health: capability-adjusted life-years, CALYs. *Appl Health Econ Health Policy*. 2017;15:437–440. doi:10.1007/s40258-017-0323-0
- Coast J, Flynn TN, Natarajan L, et al. Valuing the ICECAP capability index for older people. *Soc Sci Med*. 2008;67:874–882. doi:10.1016/j.socscimed.2008.05.015
- Husbands S, Mitchell PM, Floredin I, et al. The children and young people quality of life study: a protocol for the qualitative development of attributes for capability wellbeing measures for use in health economic evaluation with children and young people. *Wellcome Open Res*. 2022;7:117. doi:10.12688/wellcomeopenres.17801.1
- Al-Janabi H, Peters TJ, Brazier J, et al. An investigation of the construct validity of the ICECAP-A capability measure. *Qual Life Res*. 2013;22:1831–1840. doi:10.1007/s11136-012-0293-5
- Keeley T, Al-Janabi H, Lorgelly P, Coast J. A qualitative assessment of the content validity of the ICECAP-A and EQ-5D-5L and their appropriateness for use in health research. *PLoS One*. 2013;8:e85287. doi:10.1371/journal.pone.0085287
- Goranitis I, Coast J, Al-Janabi H, Lathe P, Roberts TE. The validity and responsiveness of the ICECAP-A capability-well-being measure in women with irritative lower urinary tract symptoms. *Qual Life Res*. 2016;25:2063–2075. doi:10.1007/s11136-015-1225-y
- Linton MJ, Mitchell PM, Al-Janabi H, et al. Comparing the German translation of the ICECAP-A capability wellbeing measure to the original English version: psychometric properties across healthy samples and seven health condition groups. *Appl Res Qual Life*. 2020;15:651–673. doi:10.1007/s11482-018-9681-5
- Clément V, Trouillet R, Blayac T, Ninot G. First results on the French version of the ICECAP-A questionnaire. *J De Gestion Et D'économie De La Santé*. 2019;5:385–398.
- Tang CX, Xiong Y, Wu HY, Xu J. Adaptation and assessments of the Chinese version of the ICECAP-A measurement. *Health Qual Life Outcomes*. 2018;16. doi:10.1186/s12955-018-0865-3

33. Rohrbach PJ, Dingemans AE, Groothuis-Oudshoorn CGM, et al. The ICEpop capability measure for adults instrument for capabilities: development of a tariff for the Dutch general population. *Value Health*. 2022;25:125–132. doi:10.1016/j.jval.2021.07.011
34. Farkas M, Huynh E, Gulacsi L, et al. Development of population tariffs for the ICECAP-A instrument for Hungary and their comparison with the UK tariffs. *Value Health*. 2021;24:1845–1852. doi:10.1016/j.jval.2021.06.011
35. Tokyo Metropolitan Government Bureau of Social Welfare and Public Health: The Dementia self-screening-checklist: Tokyo dementia navi. Tokyo Metropolitan Government Bureau of Social Welfare and Public Health,; 2017.
36. Tsuchiya A, Ikeda S, Ikegami N, et al. Estimating an EQ-5D population value set: the case of Japan. *Health Econ*. 2002;11:341–353. doi:10.1002/hec.673
37. Shirowa T, Ikeda S, Noto S, et al. Comparison of VALUE SET BASED on DCE and/or TTO data: scoring for EQ-5D-5L HEALTH STATES in Japan. *Value Health*. 2016;19:648–654. doi:10.1016/j.jval.2016.03.1834
38. Diener E, Emmons RA, Larsen RJ, Griffin S. The satisfaction with life scale. *J Pers Assess*. 1985;49:71–75. doi:10.1207/s15327752jpa4901_13
39. Kadono T. Development and validation of the Japanese version of the satisfaction with life scale. *Japanese Assoc Educ Psychol*. 1994;36:192.
40. Diener E, Tov W, Kim-Prieto C, Choi D, Oishi S, Biswas-Diener R. New well-being measures: short scales to assess flourishing and positive and negative feelings. *Soc Indic Res*. 2010;97:143–156. doi:10.1007/s11205-009-9493-y
41. Sumi K. Reliability and validity of Japanese versions of the flourishing scale and the scale of positive and negative experience. *Soc Indic Res*. 2013;118:601–615. doi:10.1007/s11205-013-0432-6
42. Cohen S, Kamarck T, Mermelstein R. A global measure of perceived stress. *J Health Soc Behav*. 1983;24:385–396. doi:10.2307/2136404
43. Mimura C, Griffiths P. A Japanese version of the perceived stress scale: cross-cultural translation and equivalence assessment. *BMC Psychiatry*. 2008;8:85. doi:10.1186/1471-244X-8-85
44. Mizuho Research Technologies L. Report on survey and study project on aging society and low birthrate (health awareness survey section). Edited by Ministry of Health Labour and Welfare of Japan. Tokyo, Ministry of Health Labour and Welfare of Japan. 2014;8.
45. Aizaki H, Fogarty J. An R package and tutorial for case 2 best-worst scaling. *J Choice Modelling*. 2019;32: 100171.
46. Croissant Y. Estimation of random utility models in R: the mlogit package. *J Stat Software*. 2020;95:1–14. doi:10.18637/jss.v095.i11
47. Sarrias M, Daziano RA. Multinomial logit models with continuous and discrete individual heterogeneity in R: the gmn1 package. *J Stat Software*. 2017;79:1–46. doi:10.18637/jss.v079.i02
48. Sjoberg DD, Whiting K, Curry M, Lavery JA, Larmarange J. Reproducible summary tables with the gtsummary package. *R J*. 2021;13:570–580.
49. Statistics of Bureau of Japan: Population Estimates: statistics of Japan 2022. Tokyo, Statistics of Bureau of Japan,; 2023.
50. Statistics of Bureau of Japan: Population Census 2020 Population Census Basic Complete Tabulation on Labour Force. Tokyo, Statistics of Bureau of Japan,; 2022.
51. Clancy PM. *The Acquisition of Communicative Style in Japanese*. in *Language Socialization Across Cultures*. Cambridge: Cambridge University Press; 1986:213–250. Edited by Schieffelin BB, Ochs E.

Psychology Research and Behavior Management

Publish your work in this journal

Psychology Research and Behavior Management is an international, peer-reviewed, open access journal focusing on the science of psychology and its application in behavior management to develop improved outcomes in the clinical, educational, sports and business arenas. Specific topics covered in the journal include: Neuroscience, memory and decision making; Behavior modification and management; Clinical applications; Business and sports performance management; Social and developmental studies; Animal studies. The manuscript management system is completely online and includes a very quick and fair peer-review system, which is all easy to use. Visit <http://www.dovepress.com/testimonials.php> to read real quotes from published authors.

Submit your manuscript here: <https://www.dovepress.com/psychology-research-and-behavior-management-journal>

Dovepress
Taylor & Francis Group