自己評価データに基づく can-do尺度構成法の改善

—未経験バイアスの定量的評価と補正—

独立行政法人国際交流基金
日本語試験センター
川端一光

言語教育分野における
Can-Do 尺度の概要

- Can-do-statements
  具体的な言語行動場面を記述した短い文章のこと
  cf.「英字新聞の経済欄を読むことができる」

- 言語教育分野における近年のCDS研究の発端は，CEFRの成功に依るところが大きい
- 「CEFR*」は欧州で開発された言語能力参照枠
- 特定の言語に依らない形で6段階の能力水準をCDSで記述
  → 異なる外国語間で能力基準のフレームワークを統合させられる
  → 言語教育シラバスをCEFRのレベルに準拠させることで，能力水準の統一的解釈が可能

※Common European Framework of References for Language: Learning, teaching, assessment

- 1990年代以降，大規模言語試験におけるテストスコア解釈の為に，
- テストスコアとCan-do尺度の対応づけがなされることが一般的となった。

→CEFRのレベルと言語テストのスコアをリンキングすることで，CEFR準拠のIRTスコアの解釈基準が得られるようになる。

自己評価can-do尺度

- 言語学習者に対する客観評定(教師評定)で作成されるCan-do尺度

cf. CEFRのcan-do尺度

→多相ラッシュモデルによる尺度化

- 言語学習者の自己評定で作成されるCan-do尺度

TOEIC Can-do-statements

ALTE Can-do-statements

ジェトロビジネス日本語能力テスト Can-do リポート

自己評価データによるCan-do尺度構成の問題

- 自己評価に起因する各種バイアスが回答データに混入
  → 想定されるバイアスの例 ―文化差，経験, パーソナリティ

- 特に，質問されている言語行動に関する受験者の過去の経験は，回答の信頼性・妥当性に強く影響することが予想される。実際，根岸(2006)では，経験に基づかない自己評価の精度は高くないことが報告されている。

- しかし，データ収集状況によっては，未経験CDSに対する自己評価(例えば想像して回答させたもの)の比率が大きくなってしまい，これを削除すると，利用できるデータが少数に限られてしまう場合もある。
  →海外の日本語学習者からデータを取るような場合

- 回答に混入する未経験バイアスを定量的に評価し補正することで，少なくともデータの信頼性を確保した状態で，CDSの難易度，識別力を推定することはできないか？→データの活用

本研究の目的

- 未経験項目(または経験項目)での，回答バイアス(経験効果)を評価するとともに，項目難易度，識別力，能力母数への影響を統制する項目反応モデルの提案
- 母数推定法の提案と，シミュレーション研究による精度評価(GPCMとの比較)
- 実際のCDSデータ(日本語能力試験can-do statemetns 試行版)への適用例も同時に示す

•モデリングの方針

1. CDSの評定は評定尺度で実施されることが多いから，拡張の容易な一般化部分採点モデル(Muraki, 1992)を基礎モデルとして採用する。
2. 全項目経験群と全項目未経験群にデータを分割するのではなく，項目毎に個人の経験群，未経験群への割り当てが異なることを許容する.
3. 受験者の能力水準によってバイアスの効果が変動すると仮定する
4. 推定の安定性の為にバイアスは項目間で一定とする

GPCM

$$P_{jk}(\theta_i) = \frac{\exp\left[k\alpha_j\left(\theta_i - \sum_{m=0}^{k} \beta_{jm}\right)\right]}{\sum_{l=0}^{K-1} \exp\left[l\,\alpha_j\,\left(\theta_i - \sum_{m=0}^{l} \beta_{jm}\right)\right]}$$

i = 受験者, j = 項目, k = カテゴリ
N = 受験者数, n = 項目数, K= カテゴリ総数
α_j = 識別力母数($\alpha_0 = 0$)
β_{jk} = カテゴリkの累加的閾値
θ_i = 受験者iの能力母数
$P_{jk}(\theta_i)$ = 受験者iの項目jカテゴリkへの反応確率

経験効果の評価･補正の為の項目反応モデル

$$P_{jk}(\theta_i) = \frac{\exp[(\alpha_j + \gamma_i)(k(\theta_i + \delta_i) - \sum_{m=0}^{k} \beta_{jm})]}{\sum_{l=0}^{K-1} \exp[(\alpha_j + \gamma_i)(l(\theta_i + \delta_i) - \sum_{m=0}^{l} \beta_{jm})]}$$

$$\gamma_i = \varepsilon_g \, x_{ij} y_{ig} - \varepsilon_g (1 - x_{ij}) y_{ig} \quad \cdot\cdot ①$$

$$\delta_i = \zeta_g \, x_{ij} y_{ig} - \zeta_g (1 - x_{ij}) y_{ig} \quad \cdot\cdot ②$$

x_{ij} : 受験者iの項目jの経験有無→経験あり=1, 経験なし=0

y_{ig} :レベルgの所属フラグ → 所属=1,無所属=0

ε_g : レベルgに所属することによる識別力への経験効果

ζ_g :レベルgに所属することによるステップ母数への経験効果

①と②の母数化はモデルの識別に配慮したもの

母数推定法

- MCMC(Markov Chain Monte Carlo)法によるベイズ推定
- Metropolis-Hasting アルゴリズムを採用（αのサンプリングについては(宇佐美)2010を参照した）

尤度と記法

$$L(\boldsymbol{\alpha}, \boldsymbol{\beta}1, \boldsymbol{\beta}2, \boldsymbol{\beta}3, \boldsymbol{\beta}4, \boldsymbol{\varepsilon}, \boldsymbol{\zeta}, \boldsymbol{\theta} \mid \mathbf{U}, \mathbf{X}, \mathbf{Y}) = \prod_{i=1}^{N} \prod_{j=1}^{n} \prod_{k=1}^{K} P_{jk}^{d_{ijk}}(\theta_i, x_{ij}, y_{ig})$$

$\boldsymbol{\alpha}$: n個の識別力母数ベクトル

$\boldsymbol{\beta}1, \boldsymbol{\beta}2, \boldsymbol{\beta}3, \boldsymbol{\beta}4$: n個のステップ母数ベクトル

$\boldsymbol{\varepsilon}, \boldsymbol{\zeta}$: 経験効果パラメタベクトル

$\boldsymbol{\theta}$: N個の能力母数ベクトル

$\mathbf{U}$: $N \times n$の項目反応行列 (多値型項目反応)

$\mathbf{X}$: $N \times n$の項目経験行列 (2値型名義変数)

$\mathbf{Y}$: $N \times$ Gのレベル行列 (2値型名義変数)

d_{ijk} : 受験者 iの項目 jのカテゴリ kへの反応を示す 2値名義変数

事後分布・事前分布・候補分布

条件付き事後分布

$$P(\boldsymbol{\alpha} \mid \mathbf{U}, \mathbf{X}, \mathbf{Y}) \propto L(\boldsymbol{\alpha}, \boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \boldsymbol{\beta}_3, \boldsymbol{\beta}_4, \boldsymbol{\varepsilon}, \boldsymbol{\zeta}, \boldsymbol{\theta} \mid \mathbf{I}, \mathbf{X}) P(\boldsymbol{\alpha})$$

$$P(\boldsymbol{\beta}_k \mid \mathbf{U}, \mathbf{X}, \mathbf{Y}) \propto L(\boldsymbol{\alpha}, \boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \boldsymbol{\beta}_3, \boldsymbol{\beta}_4, \boldsymbol{\varepsilon}, \boldsymbol{\zeta}, \boldsymbol{\theta} \mid \mathbf{U}, \mathbf{X}, \mathbf{Y}) P(\boldsymbol{\beta}_k)$$

$$P(\boldsymbol{\varepsilon}, \boldsymbol{\zeta} \mid \mathbf{U}, \mathbf{X}, \mathbf{Y}) \propto L(\boldsymbol{\alpha}, \boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \boldsymbol{\beta}_3, \boldsymbol{\beta}_4, \boldsymbol{\varepsilon}, \boldsymbol{\zeta}, \boldsymbol{\theta} \mid \mathbf{U}, \mathbf{X}, \mathbf{Y}) P(\boldsymbol{\varepsilon}) P(\boldsymbol{\zeta})$$

$$P(\theta_i \mid \mathbf{U}_i, \mathbf{X}_i, \mathbf{Y}_i) \propto L(\boldsymbol{\alpha}, \boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \boldsymbol{\beta}_3, \boldsymbol{\beta}_4, \boldsymbol{\varepsilon}, \boldsymbol{\zeta}, \theta_i \mid \mathbf{U}_i, \mathbf{X}_i, \mathbf{Y}_i) P(\theta_i)$$

事前分布

$$\alpha_j \sim LogNormal(\mu_\alpha = 0, {\sigma_\alpha}^2 = 0.5)$$

$$\beta_{jk} \sim Normal(\mu_\beta = 0, {\sigma_\beta}^2 = 1)$$

$$\varepsilon \sim Normal(\mu_\varepsilon = 0, {\sigma_\varepsilon}^2 = 1)$$

$$\zeta \sim Normal(\mu_\zeta = 0, {\sigma_\zeta}^2 = 1)$$

$$\theta_i \sim Normal(\mu_{\theta_i} = 0, {\sigma_{\theta_i}}^2 = 1)$$

候補分布

$$\alpha^*{}_j \sim Normal(\mu = \alpha^t{}_j, \sigma^2 = 0.05)$$

$$\beta^*{}_{jk} \sim Normal(\mu = \beta^t{}_{jk}, \sigma^2 = 0.05)$$

$$\varepsilon^* \sim Normal(\mu = \varepsilon^t, \sigma^2 = 0.05)$$

$$\zeta^* \sim Normal(\mu = \zeta^t, \sigma^2 = 0.05)$$

$$\theta^*{}_i \sim Normal(\mu = \theta^t{}_i, \sigma^2 = 0.05)$$

受容率αとサンプリングの順序

識別力母数→ステップ母数→経験効果母数→能力母数

$$\alpha_{\boldsymbol{\alpha}} = \min\left\{ \frac{L(\boldsymbol{\alpha}^*, \boldsymbol{\beta}^t{}_1, \boldsymbol{\beta}^t{}_2, \boldsymbol{\beta}^t{}_3, \boldsymbol{\varepsilon}^t, \zeta^t, \boldsymbol{\theta}^t \mid \mathbf{U}, \mathbf{X}, \mathbf{Y}) P(\boldsymbol{\alpha}^*)}{L(\boldsymbol{\alpha}^t, \boldsymbol{\beta}^t{}_1, \boldsymbol{\beta}^t{}_2, \boldsymbol{\beta}^t{}_3, \boldsymbol{\varepsilon}^t, \zeta^t, \boldsymbol{\theta}^t \mid \mathbf{U}, \mathbf{X}, \mathbf{Y}) P(\boldsymbol{\alpha}^t)}, 1 \right\}$$

$$\alpha_{\boldsymbol{\beta}_1} = \min\left\{ \frac{L(\boldsymbol{\alpha}^{t+1}, \boldsymbol{\beta}^*{}_1, \boldsymbol{\beta}^t{}_2, \boldsymbol{\beta}^t{}_3, \boldsymbol{\varepsilon}^t, \zeta^t, \boldsymbol{\theta}^t \mid \mathbf{U}, \mathbf{X}, \mathbf{Y}) P(\boldsymbol{\beta}_1{}^*)}{L(\boldsymbol{\alpha}^{t+1}, \boldsymbol{\beta}^t{}_1, \boldsymbol{\beta}^t{}_2, \boldsymbol{\beta}^t{}_3, \boldsymbol{\varepsilon}^t, \zeta^t, \boldsymbol{\theta}^t \mid \mathbf{U}, \mathbf{X}, \mathbf{Y}) P(\boldsymbol{\beta}_1{}^t)}, 1 \right\}$$

$$\alpha_{\boldsymbol{\varepsilon}, \boldsymbol{\zeta}} = \min\left\{ \frac{L(\boldsymbol{\alpha}^{t+1}, \boldsymbol{\beta}^{t+1}{}_1, \boldsymbol{\beta}^{t+1}{}_2, \boldsymbol{\beta}^{t+1}{}_3, \boldsymbol{\varepsilon}^*, \boldsymbol{\zeta}^*, \boldsymbol{\theta}^t \mid \mathbf{U}, \mathbf{X}, \mathbf{Y}) P(\boldsymbol{\varepsilon}^*) P(\boldsymbol{\zeta}^*)}{L(\boldsymbol{\alpha}^{t+1}, \boldsymbol{\beta}^{t+1}{}_1, \boldsymbol{\beta}^{t+1}{}_2, \boldsymbol{\beta}^{t+1}{}_3, \boldsymbol{\varepsilon}^t, \boldsymbol{\zeta}^t, \boldsymbol{\theta}^t \mid \mathbf{U}, \mathbf{X}, \mathbf{Y}) P(\boldsymbol{\varepsilon}^t) P(\boldsymbol{\zeta}^t)}, 1 \right\}$$

$$\alpha_{\theta_i} = \min\left\{ \frac{L(\boldsymbol{\alpha}^{t+1}, \boldsymbol{\beta}^{t+1}{}_1, \boldsymbol{\beta}^{t+1}{}_2, \boldsymbol{\beta}^{t+1}{}_3, \boldsymbol{\varepsilon}^{t+1}, \boldsymbol{\zeta}^{t+1}, \theta_i{}^* \mid \mathbf{U}, \mathbf{X}, \mathbf{Y}) P(\theta_i{}^*)}{L(\boldsymbol{\alpha}^{t+1}, \boldsymbol{\beta}^{t+1}{}_1, \boldsymbol{\beta}^{t+1}{}_2, \boldsymbol{\beta}^{t+1}{}_3, \boldsymbol{\varepsilon}^{t+1}, \boldsymbol{\zeta}^{t+1}, \theta_i{}^t \mid \mathbf{U}, \mathbf{X}, \mathbf{Y}) P(\theta_i{}^t)}, 1 \right\}$$

★ステップ母数はβ1からβ4まで逐次的にサンプリングする(上記はβ1のみ)
★正規分布の対称性から候補分布は受容率の計算に含まれない

シミュレーション研究

- 項目数J=10
- 受験者数N＝1000，500，10
- 受験レベル数G=3
- 各条件でデータセットを20生成(R=20)
- カテゴリ数=5

母数真値発生状況

- α ~ Uniform[0.7, 1.2]
- β1 ~Uniform[-1.0,-0.6], β2 ~Uniform[-0.5, 1.0]
- β3 =-beta2, β4 = -beta1
- ε ~ Uniform[-0.5,0.5], ζ ~ Uniform[-1.0,1.0]

能力と経験行列の生成

- 能力と経験状況には相関があると考えるのが自然→ここでは0.5を設定
- CDSが測定するθ_1と，各項目の経験フラグを説明するθ_2を定義
- Θ2は経験度

$$[\theta_1\ \theta_2] \sim \mathrm{MVN}\ (\mu = [0,0],\ \rho = 0.5)$$

- θ_1と項目母数の真値を利用して多値カテゴリ反応行列Uを生成
- 位置母数のみを含んだロジスティック関数を利用して，経験行列Xを生成

$$P_{jA}(\theta_{2i}) = \frac{\exp(-\theta_{2i} + b_j)}{1 + \exp(-\theta_{2i} + b_j)} \qquad X_i = \begin{cases} x_{ij} = 1,\ P_{jA}(\theta_{2i}) > U[0,1] \\ x_{ij} = 0,\ else \end{cases}$$

bjは項目jの経験のし難さを表現
bj ~ Normal(0,1) とする

レベル行列の生成
-受験者iの場合を例に-

$$y_{i1} = \begin{cases} y_{i1} = 1, \theta_{1i} > 0.5 \\ y_{i1} = 0, \quad\quad else \end{cases}$$

$$y_{i2} = \begin{cases} y_{i2} = 1, -0.5 < \theta_{1i} \leq 0.5 \\ y_{i2} = 0, \quad\quad else \end{cases}$$

$$y_{i3} = \begin{cases} y_{i3} = 1, \theta_{1i} \leq -0.5 \\ y_{i3} = 0, \quad\quad else \end{cases}$$

$$Y_i = [y_{i1}\ y_{i2}\ y_{i3}]$$

サンプリング回数と収束判定

- 全シミュレーション条件で90000回の更新
- Burn-in期間は1：40000
- 一つの長いマルコフ連鎖を生成し，Gewekeの指標によって，事後分布への収束を確認した。

- Gewekeの指標の観点では連鎖の収束が示唆されない母数もあったが，サンプリングの履歴がホワイトノイズの様相を示しており，事後分布の標準偏差が十分に小さく，かつ真値を適切に回復していることから，収束していると判断した。

精度比較

- 各条件の20のデータセットに対して，それぞれ母数推定を行なう
- マルコフ連鎖が発散している母数が存在した場合には，そのデータセットの結果は破棄する
- 上述のシミュレーションデータに対して，提案モデルとGPCMを同時に適用し，経験効果に対する提案モデルの耐性と，経験効果の検出力を評価する

x:推定値，t:真値，r:データセットID

BIAS＝ $[\sum_{r=1}^{R} x_r - t_r]/R$

RMS＝ $\sqrt{[\sum_{r=1}^{R}(x_r - t_r)^2]/R}$

シミュレーション結果

効果パラメタの推定精度

N=1000, J=10

	ε	ζ
BIAS	0.052	0.034
RMS	0.166	0.151

N=500, J=10

	ε	ζ
BIAS	0.026	0.023
RMS	0.177	0.167

N=100, J=10

	ε	ζ
BIAS	0.032	0.042
RMS	0.195	0.178

N=1000, J=10

		α	β1	β2	β3	β4	θ
BIAS	提案法	0.099	0.081	0.045	0.044	0.077	0.074
	GPCM	0.221	0.098	0.175	0.27	0.142	0.082
RMS	提案法	0.128	0.185	0.187	0.183	0.182	0.364
	GPCM	0.272	0.336	0.35	0.462	0.271	0.410
ratio	BIAS	2.222	1.21	3.894	6.196	1.847	1.108
	RMS	2.128	1.816	1.878	2.53	1.487	1.126

Ratio= GPCM / 提案法

提案モデルの推定精度が一貫して高い

N=500, J=10

		α	β1	β2	β3	β4	θ
BIAS	提案法	0.1	0.064	0.068	0.059	0.057	0.076
	GPCM	0.233	0.164	0.254	0.275	0.171	0.088
RMS	提案法	0.134	0.221	0.262	0.218	0.185	0.380
	GPCM	0.289	0.392	0.564	0.468	0.372	0.434
ratio	BIAS	2.327	2.544	3.759	4.670	3.012	1.158
	RMS	2.155	1.776	2.156	2.144	2.014	1.142

N=100,J=10

		α	β1	β2	β3	β4	θ
BIAS	提案法	0.095	0.113	0.136	0.088	0.145	0.094
	GPCM	0.118	0.117	0.192	0.240	0.194	0.103
RMS	提案法	0.194	0.372	0.374	0.354	0.390	0.384
	GPCM	0.281	0.522	0.677	0.650	0.558	0.419
ratio	BIAS	1.239	1.037	1.410	2.729	1.331	1.092
	RMS	1.448	1.403	1.812	1.836	1.430	1.090

シミュレーション研究考察

- シミュレーションの結果，GPCMに対する提案モデルの母数回復精度の高さが確認された。
- 項目に関する経験は容易に測定できるので，自己評定によるcan-do尺度構成法の分析モデルとして提案モデルは有望。実務レベルでの適用可能性は極めて高い。
- シミュレーションの条件が十分でない。少なくとも項目数を変化させる必要がある

適用例

JLPT-CDS試行版

(国際交流基金，財団法人日本国際支援協会)

- CEFR, ACTFL-OPI, TOEIC can-do statementsをもとに作成した日本語能力のCDS(長沼他，2007;野口他，2006)
- 「聞く(L)」「読む(R)」「話す(S)」「書く(W)」の４技能をそれぞれ20項目で測定
- 各項目の過去の経験の有無を問う

If 経験あり

「全然できなかった」 ←12345→ 「問題なくできた」

else if 経験なし

「全然できないと思う」←12345→ 「問題なくできると思う」

試行版作成の履歴

- 2005年11月～12月
  学習者(135名)に対する予備調査
  専門家(30名)に対する予備調査
- 2006年1月
  CDS文言確定→翻訳(英・中・韓・インドネシア語)
- 2006年3月本調査(1068名)

2006年3月本調査

- 2005年12月日本語能力試験受験者からモニタを募集(上位[1級，2級，3級，4級]下位)

- 中国語母語話者　393名
- 韓国語母語話者　132名
- インドネシア語母語話者　127名
- その他　416名
- 合計　1068名

分析データ

- １級，２級，３級の欠測のない受験者を抽出

- 1級：157名
- 2級：334名
- 3級：201名

一因子性の確認

	L	R	S	W
α係数	0.934	0.959	0.941	0.950
因子負荷量 二乗和	8.482	10.816	9.036	9.896
寄与率	0.424	0.541	0.452	0.495
第１固有値寄与	0.452	0.564	0.479	0.520
第２固有値寄与	0.064	0.070	0.074	0.091

未経験回答数

	L			R		
項目	1級	2級	3級	1級	2級	3級
項目1	44	114	82	2	12	23
項目2	1	3	19	14	35	56
項目3	4	17	28	26	92	67
項目4	2	8	26	2	13	40
項目5	20	71	58	1	13	37
項目6	2	6	36	7	40	73
項目7	9	38	48	12	46	66
項目8	21	59	78	2	22	9
項目9	41	114	97	12	33	22
項目10	61	161	59	11	32	42
項目11	0	0	8	3	33	79
項目12	3	7	21	16	74	94
項目13	16	39	41	22	110	114
項目14	11	46	52	29	138	105
項目15	4	16	18	14	45	37
項目16	10	37	29	30	115	105
項目17	13	49	43	60	201	138
項目18	54	165	87	4	36	36
項目19	0	5	11	2	18	16
項目20	6	29	26	24	94	64
合計	322	984	867	293	1202	1223

未経験と回答した度数が少ない項目が存在する

未経験回答数

	S			W		
項目	1級	2級	3級	1級	2級	3級
項目1	0	1	2	5	25	32
項目2	7	29	22	30	75	72
項目3	11	38	42	40	88	105
項目4	7	19	26	26	52	71
項目5	31	81	68	72	160	125
項目6	0	4	6	53	99	78
項目7	6	17	23	15	45	37
項目8	15	90	78	32	99	95
項目9	91	247	121	81	199	147
項目10	52	152	98	77	190	152
項目11	10	20	11	0	1	11
項目12	6	6	19	16	62	111
項目13	25	91	63	24	126	113
項目14	44	98	70	15	36	46
項目15	12	37	31	40	129	94
項目16	4	20	11	5	24	26
項目17	5	17	16	30	88	59
項目18	27	88	67	109	272	148
項目19	38	122	90	81	212	162
項目20	54	163	98	78	245	151
合計	445	1340	962	829	2227	1835

推定結果

- ４技能についてそれぞれ個別に分析を実行
- MCMC反復を90000回とし，最初の40000回をBurn-in期間とした。
- 「W」については，90000回で多数の母数のマルコフ連鎖が事後分布に収束しなかったため，180000回までMCMC反復を増やした
- 事後分布の標準偏差を確認し，全パラメタの推定精度が高いことを確認した。

母数推定値

	L	R	S	W
ε1	-0.398	-0.376	-0.354	-0.196
ε2	-0.261	-0.231	-0.150	-0.091
ε3	-0.177	-0.062	-0.239	-0.322
ζ1	-0.274	-0.361	-0.469	-0.332
ζ2	-0.338	-0.343	-0.452	-0.420
ζ3	-0.192	-0.146	-0.416	-0.292

●全レベル・全技能において未経験であることによって識別力が低下する

●全レベル・全技能において未経験であることによって困難度が上昇する

MC-SE(事後分布の標準偏差)

	L	R	S	W
ε1	0.044	0.063	0.044	0.048
ε2	0.029	0.038	0.032	0.035
ε3	0.037	0.059	0.042	0.056
ζ1	0.039	0.041	0.037	0.026
ζ2	0.023	0.020	0.025	0.018
ζ3	0.025	0.020	0.027	0.023

●経験効果母数の標準誤差は十分小さい

βに対する経験効果の解釈

	L	R	S	W
ζ1	-0.274	-0.361	-0.469	-0.332
ζ2	-0.338	-0.343	-0.452	-0.420
ζ3	-0.192	-0.146	-0.416	-0.292

経験群の平均ー未経験群の平均

L	1級	2級	3級	R	1級	2級	3級
M	0.333	0.172	0.130	M	0.755	0.741	0.568
SD	-0.048	0.049	-0.085	SD	-0.122	-0.007	0.003
S	1級	2級	3級	W	1級	2級	3級
M	0.823	0.879	0.753	M	0.740	0.808	0.755
SD	-0.136	0.023	0.083	SD	0.027	0.037	0.082

★全レベルで平均値差は正になっていることに対応して，ζは全て負になっている。

αに対する経験効果の解釈

	L	R	S	W
ε1	-0.398	-0.376	-0.354	-0.196
ε2	-0.261	-0.231	-0.150	-0.091
ε3	-0.177	-0.062	-0.239	-0.322

未経験群の評定の度数

レベル	１級					２級					３級				
カテゴリ	1	2	3	4	5	1	2	3	4	5	1	2	3	4	5
L	14	74	140	67	27	112	359	418	85	19	158	293	346	51	19
R	13	53	118	77	32	97	452	448	162	43	340	392	369	77	58
S	46	117	187	70	25	186	556	453	120	45	225	330	327	44	36
W	64	211	371	154	29	350	872	730	237	38	580	642	486	91	36

L，R，S：特に1級で3付近に回答が集中　→　経験していないので、「問題なくできる」とはいいきれないが，「全くできない」というほど自信が無いわけではない
W：特に３級で１付近に回答が集中　→　「書く」タスクの難易度がそもそも高く、低レベルの受験者にとって，経験していない場合には，一様に難しくとらえられる

適用例考察

- ステップ母数だけでなく，識別力にも経験効果が存在する
- 能力レベルによって経験効果は異なる
- 基礎統計量と整合性のある推定値が得られた
- シミュレーション研究の結果から，これらの効果を統制した上で，GPCMの識別力母数とステップ母数を推定することが可能

展望

- 未経験項目の回答には，妥当性がない可能性もある。特に自己モニタリング力が低い，初学者の回答を扱う場合には，この点についての限界がある。
- 母語要因もCDSへの回答に影響を与える可能性が高いが，本研究では取り上げていない。
- 提案モデルを多母集団モデルとして表現することで対応できると考えられる。DIF研究と統合することも可能かもしれない。

文献

- 野口裕之・熊谷龍一・大隅敦子・石毛順子・長沼君主 (2006). 日本語能力試験can-do statements(試行版)のIRT尺度化と日本語能力試験の得点段階の対応づけの試み 5th International J-OPI-Symposium Berlin　2006.

- 長沼君主・大隅敦子・和田晃子・伊東祐朗・熊谷龍一・野口裕之(2007). JLPT 日本語能力記述文作成の試み–日本語能力試験(JLPT)Can-do Statements 試行版の分析から2007 年度日本語教育学会春季大会予稿集，pp215-218.

- 宇佐美慧(2010). 採点者側と受験者側のバイアス要因の影響を同時に評価する多値型項目反応モデル–MCMC アルゴリズムに基づく推定– 教育心理学研究，58, 163-175.

Ikko_Kawahashi@jpf.go.jp

※本研究は発表者の個人研究であり，国際交流基金，財団法人日本国際支援協会の活動とは一切関係ありません。