

実施データに基づく  
**VELC Test** の信頼性・妥当性  
の検証

静 哲人  
(大東文化大学)

# 実施状況

- 2012年度 試験的に無料実施
  - 32大学(高専含む) 12,366名が受験  
(2013年度より 正式に実施 @800円税抜)
- 32大学のうち、 $N > 100$ であった18校
- 100名を無作為抽出 × 18校
- この1800名のデータを分析

# 2012年度データでの信頼性検証

受験者能力値の信頼性，項目難度の信頼性に関わる指標の値

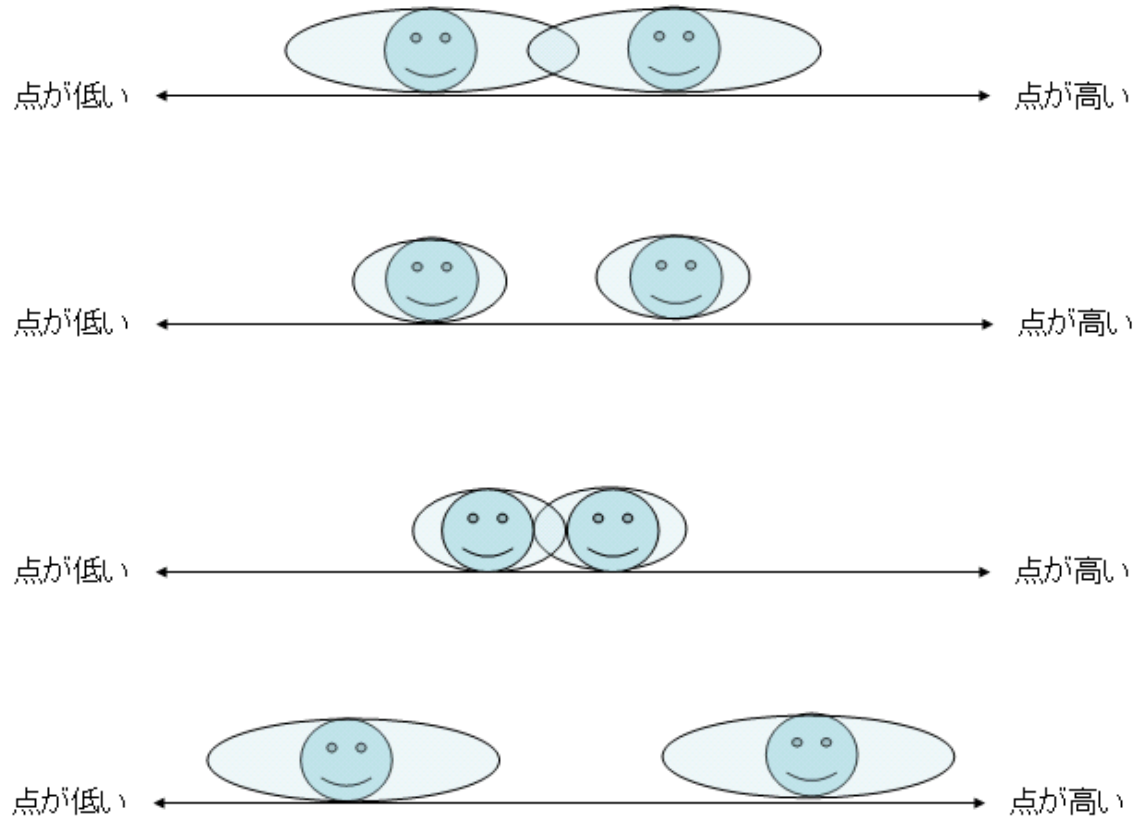
	項目数	N	PR	PS	IR	IS
TTL	120	1800	0.94	4.04	1.00	18.13
L	60	1800	0.86	2.53	1.00	18.05
R	60	1800	0.91	3.26	1.00	17.42
L1	20	1800	0.68	1.46	1.00	19.29
L2	20	1800	0.68	1.40	1.00	18.54
L3	20	1800	0.72	1.59	0.99	12.72
R1	20	1800	0.75	1.72	1.00	17.92
R2	20	1800	0.76	1.77	0.99	13.26
R3	20	1800	0.78	1.90	0.99	13.20

PR: person reliability PS: person separation IR : Item reliability IS: item separation

能力値: LよりR、特にR3で差がつきやすい傾向が見える

項目難度値: 非常に難度差幅が広い、とくにL1, L2が広いことが分かる。

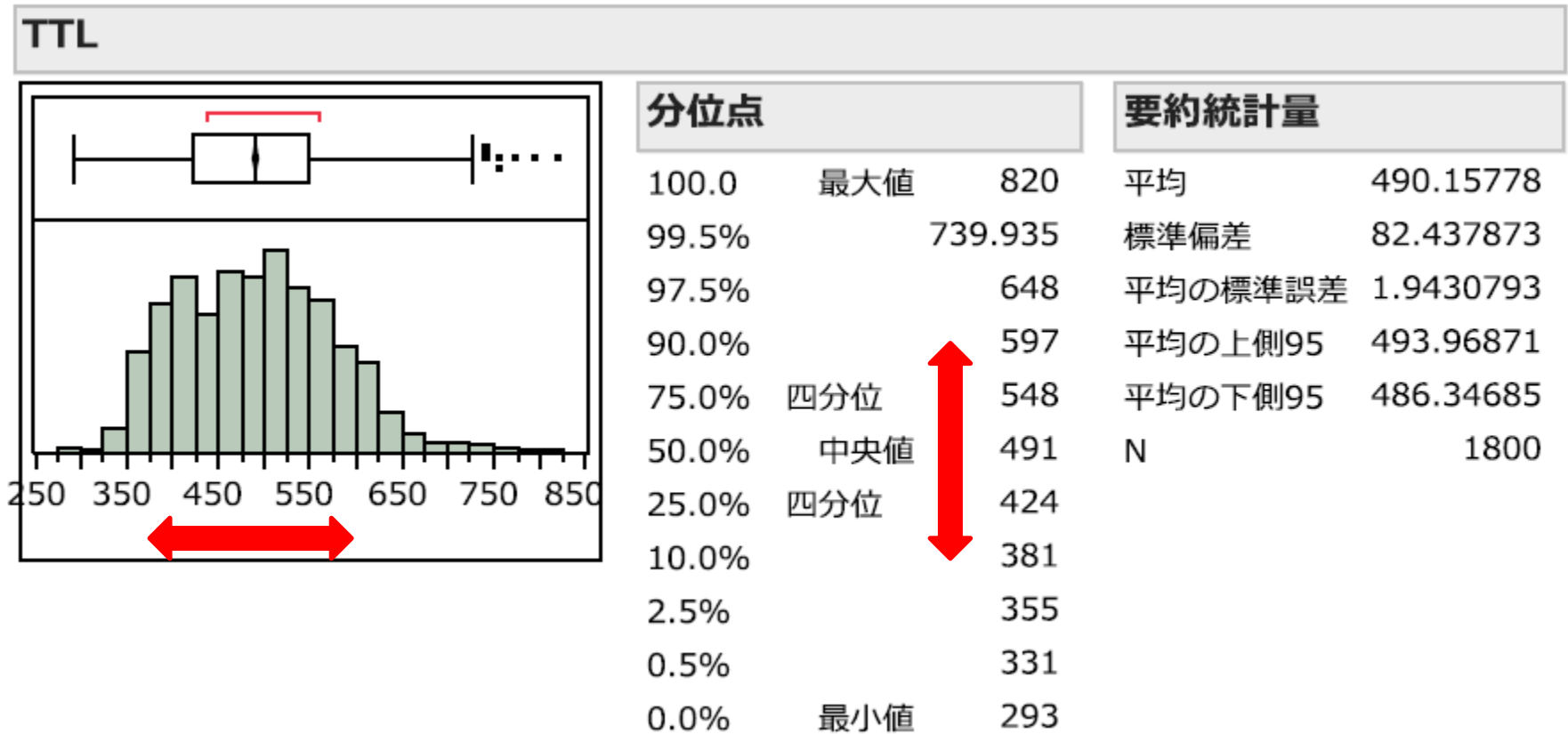
# 信頼性= 順位の変動性



# 指標の意味

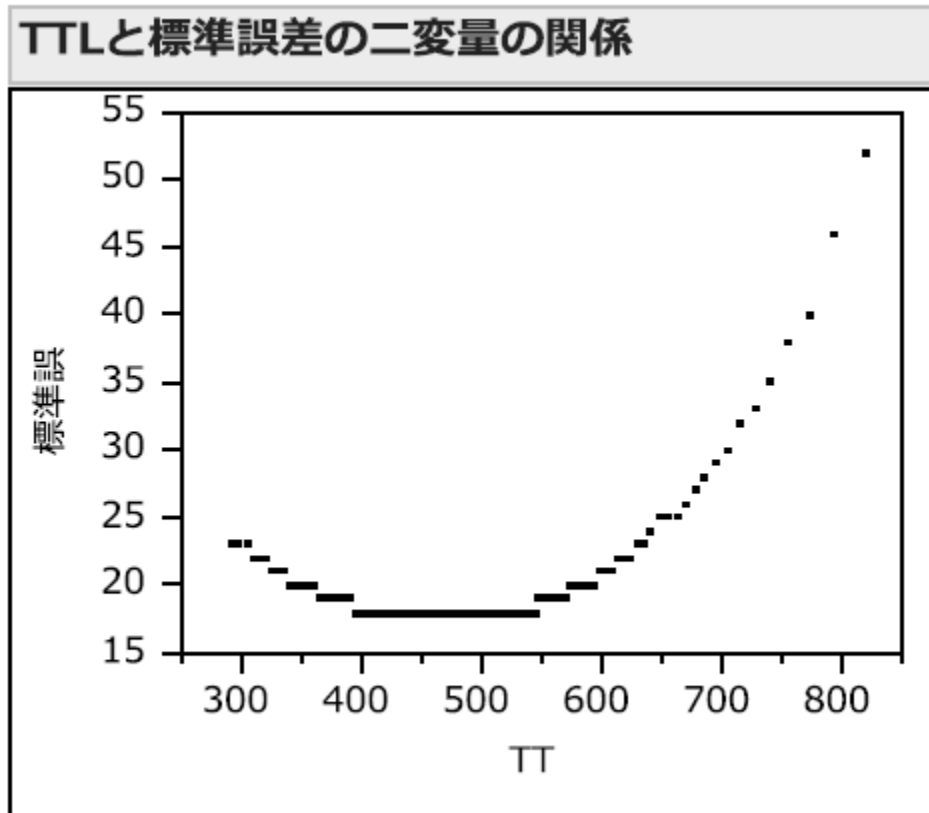
- PR = Person Reliability (上限値=1)
  - 真の能力値分散 / 推定された能力値分散
- PS = Person Separation (上限値なし)
  - 統計的に分離され得る受験者能力層の数
- IR = Item Reliability (上限値=1)
  - 真の難度値分散 / 推定された難度値分散
- IS = Item Separation (上限値なし)
  - 統計的に分離されうる項目難度層の数

# VELC総合スコアの分布と誤差分布



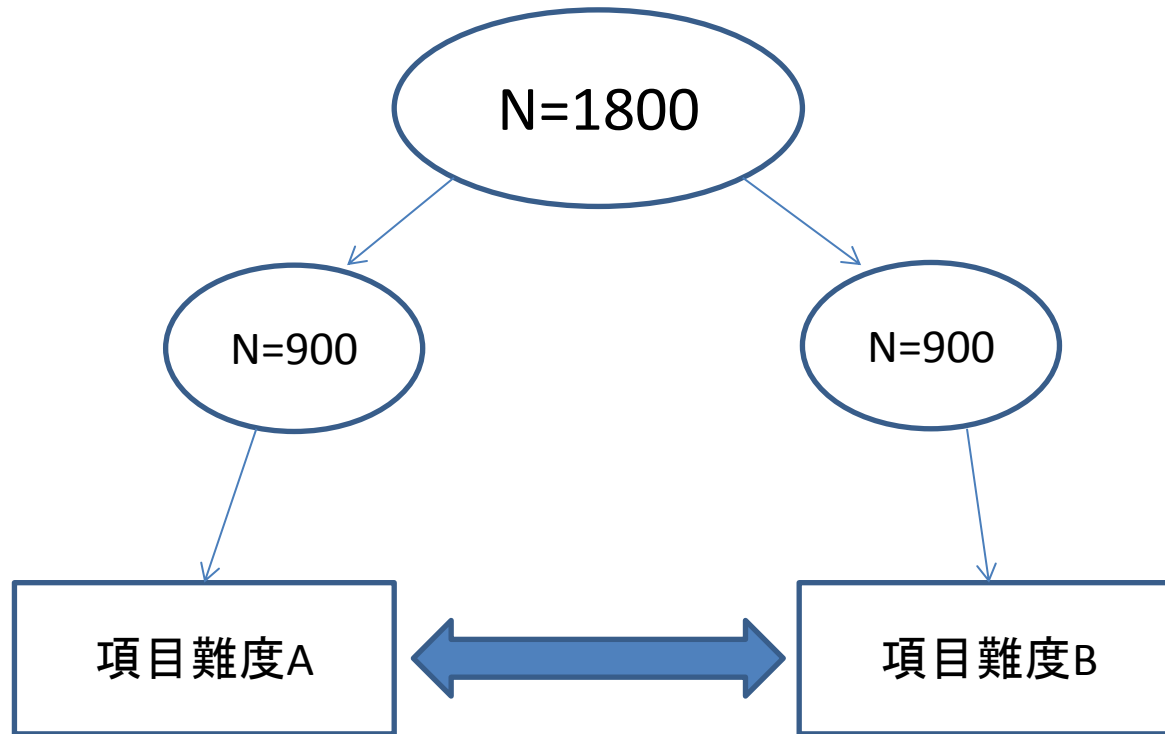
381 ~ 597 に、中核的な80%が分布している。

# 標準誤差



最も中核的な受験者層において、もっとも誤差がすくなくくなるような項目難度設定になっている。

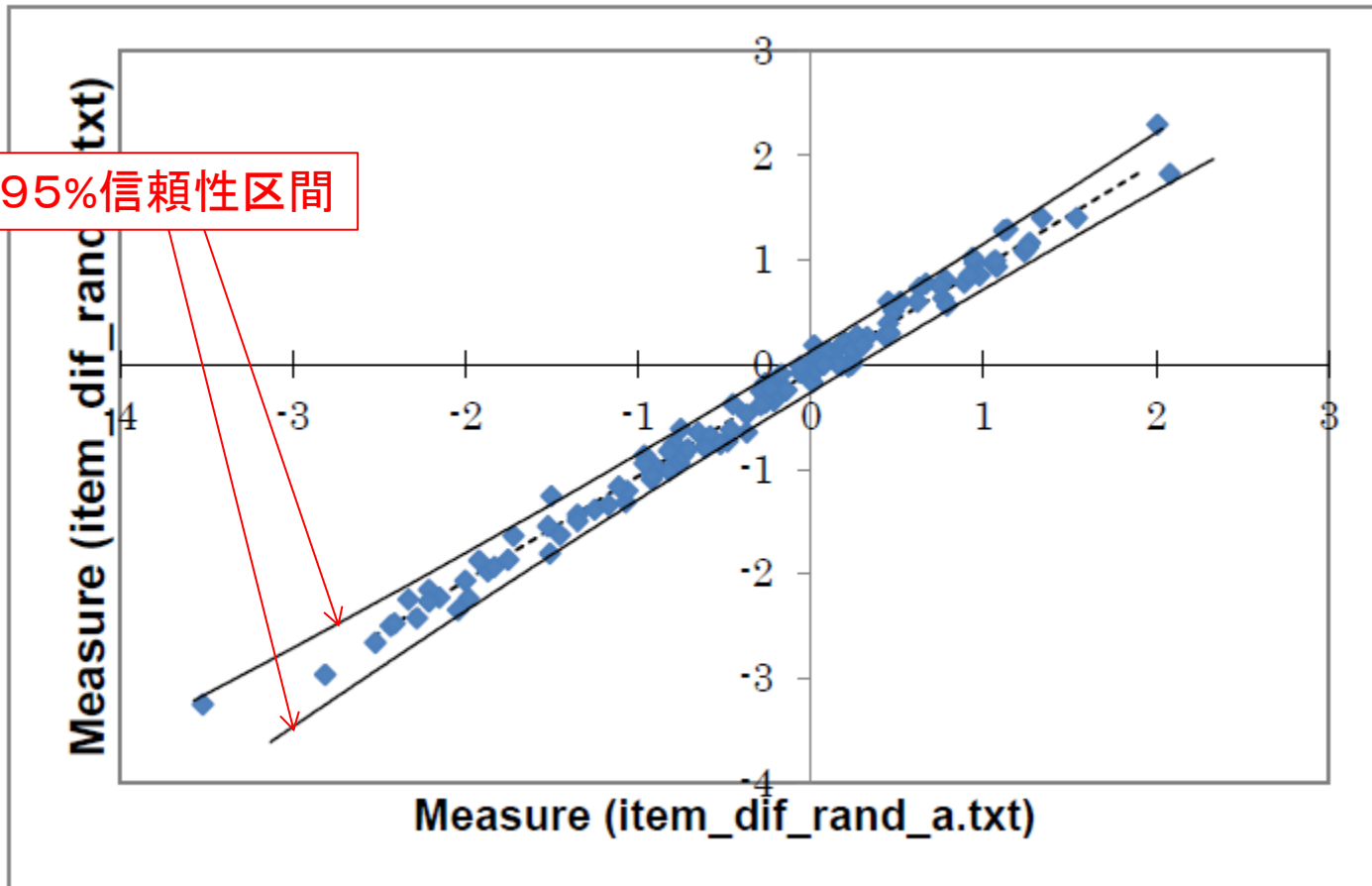
# 項目難度の安定性



ランダム分割の900名ごとに項目難度を改めて推定  
→ 2セットの項目難度の一致度を見るためプロット



# 項目難度の安定性

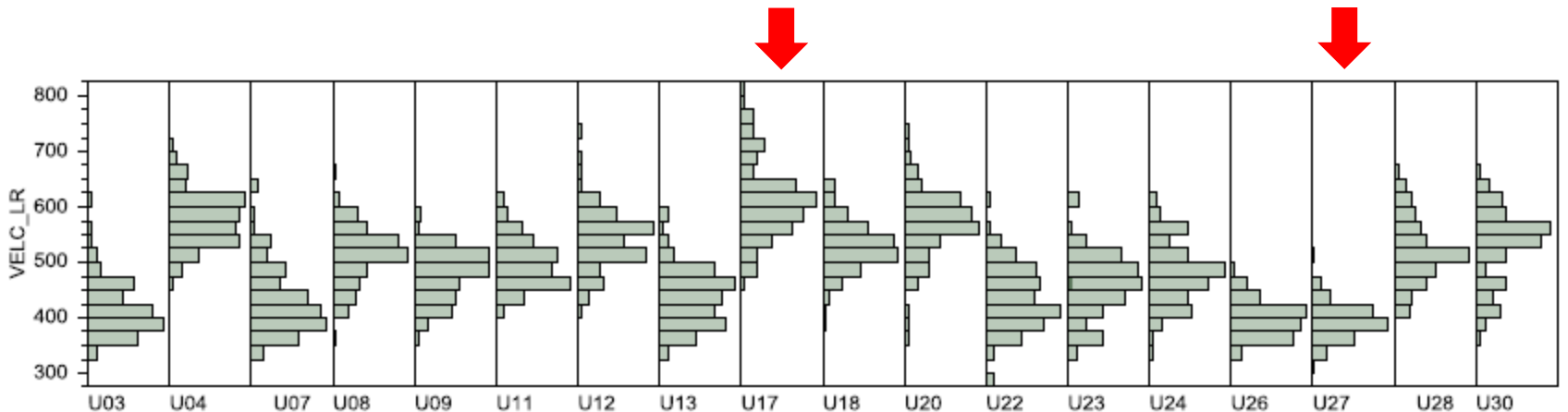


ランダム分割の900名ごとに推定した項目難度のプロット  $r = .994$   
特定の受験者サンプルに影響されずに安定している。  
実際に使用するアンカー値との相関は、いずれも  $r = .971$

# 異なる大学間に差はあるか

大学別の TTL の記述統計

	U03	U04	U07	U08	U09	U11	U12	U13	U17
Mean	415.9	577.4	430.0	513.2	480.3	500.9	542.0	437.0	612.3
SD	50.5	45.9	61.6	51.7	47.2	46.7	54.45	51.6	70.6
	U18	U20	U22	U23	U24	U26	U27	U28	U30
Mean	531.6	568.6	434.8	459.9	481.4	396.5	393.5	520.4	527.1
SD	47.8	63.5	57.7	57.2	56.9	32.1	32.6	59.1	68.5



平均値は 612.3 ~ 393.5 有意な差あり。 効果量 = 大

# 18大学間での多重比較 (Tukey-Kramer HSD 検定)の結果

Level		Mean
U17	A	612.3
U04	B	577.4
U20	B C	568.6
U12	C D	542.0
U18	D E	531.6
U30	D E F	527.1
U28	D E F	520.4
U08	E F	513.2
U11	F G	500.9
U24	G H	481.4
U09	G H	480.3
U23	H I	459.9
U13	I J	437.0
U22	I J	434.9
U07	J	430.0
U03	J K	416.0
U26	K	396.5
U27	K	393.5

同じ文字でつながっている大学は有意差がない

能力の重なりのある11グループに分かれる

# 18大学間での多重比較 (Tukey-Kramer HSD 検定)の結果

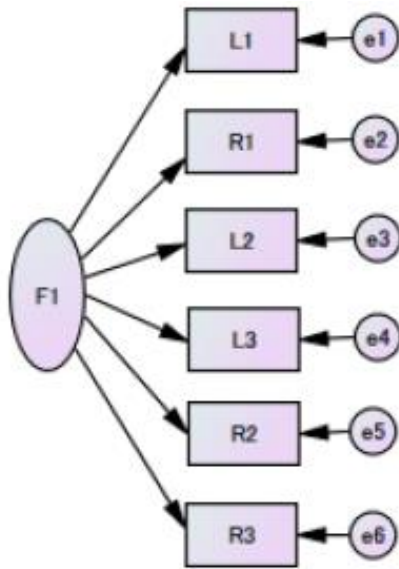
Level		Mean
U17	A	612.3
U04	B	577.4
U20	B C	568.6
U12	C D	542.0
U18	D E	531.6
U30	D E F	527.1
U28	D E F	520.4
U08	E F	513.2
U11	F G	500.9
U24	G H	481.4
U09	G H	480.3
U23	H I	459.9
U13	I J	437.0
U22	I J	434.9
U07	J	430.0
U03	J K	416.0
U26	K	396.5
U27	K	393.5

同じ文字でつながっている大学は有意差がない

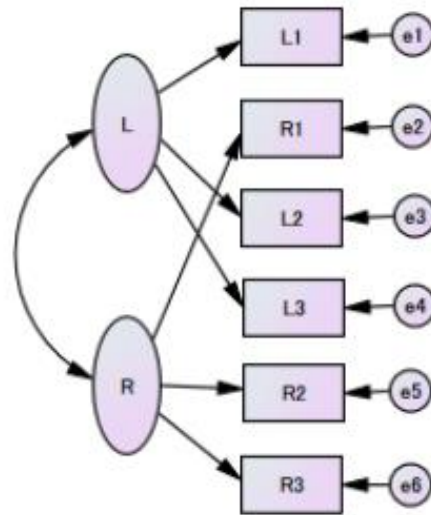
能力の重なりのない6グループに分かれる

# どのような能力を測っているのか

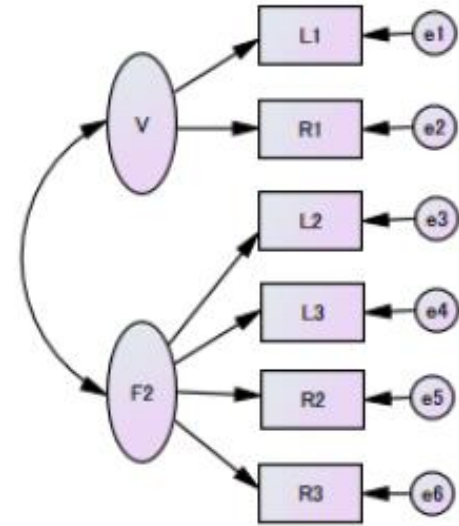
- $N = 1800$
- L1, L2, L3, R1, R2, R3 の6変数を観測変数として、その裏にある因子構造を探るため、AMOSを用いた構造方程式モデリング(SEM)を試みた。
- 理論的に想定される複数の競合モデルを当てはめ、適合度指標を比較した。



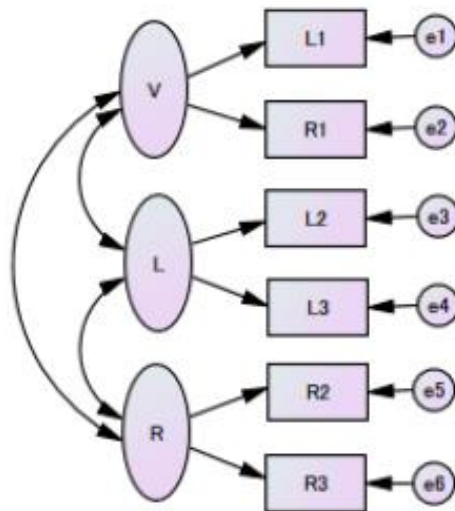
モデル 1



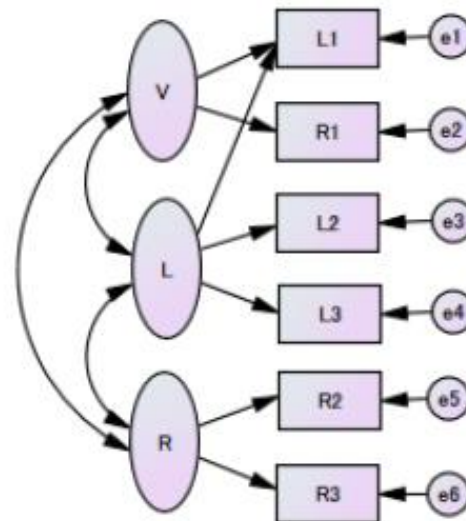
モデル 2



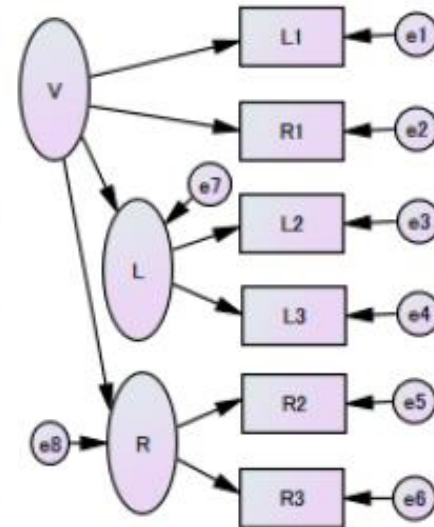
モデル 3



モデル 4



モデル 5

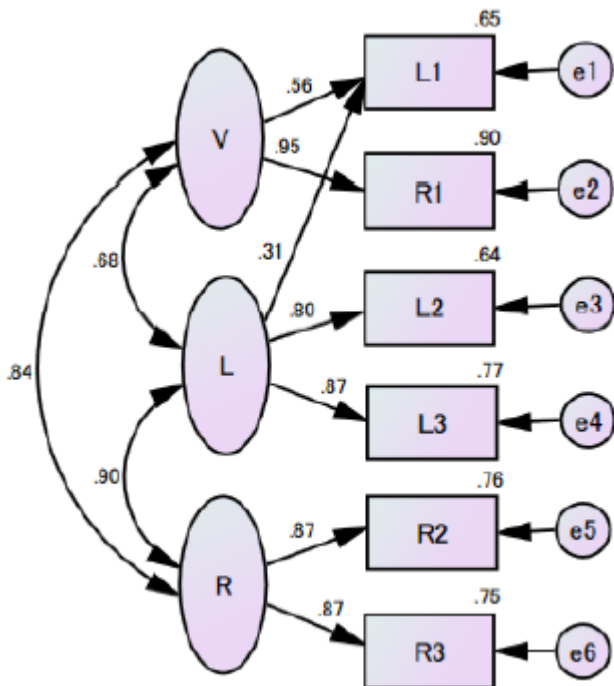


モデル 6

# 適合度指標の比較

6モデルの適合度指標

	Chi-sq.	df	p	GFI	AGFI	CFI	RMSEA	SRMR	AIC
モデル 1	435.782	9	0.000	0.923	0.820	0.942	0.162	0.042	459.782
モデル 2	403.601	8	0.000	0.926	0.805	0.947	0.166	0.042	429.601
モデル 3	223.053	8	0.000	0.957	0.888	0.971	0.122	0.030	249.053
モデル 4	58.393	6	0.000	0.990	0.963	0.993	0.070	0.014	88.393
モデル 5	7.744	5	0.171	0.999	0.994	1.000	0.017	0.005	39.744
モデル 6	241.263	7	0.000	0.958	0.874	0.968	0.136	0.031	269.263



語彙力がL1, R1に影響し、  
聴く力がL1, L2, L3に影響し、  
読む力が、R2, R3に影響する、  
と解釈。

文脈なしの語彙問題を設定して  
いるベルクテストの特性をとらえた  
適切なモデル、と考えられる。

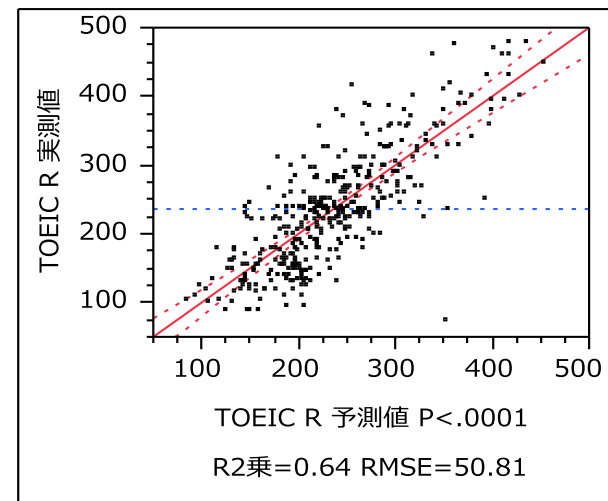
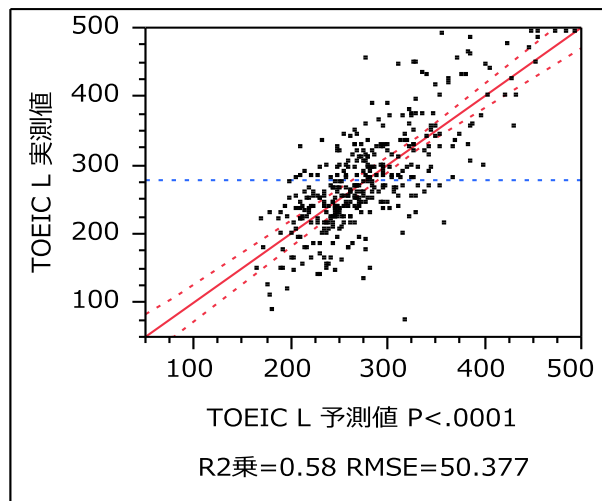
# まとめ

- 本テストは日本人大学生の平均的層に適切に  
的を絞った難度設定を行ったことにより、信頼性  
が高い。[信頼性を示唆]
- 項目難度は受験者が変わっても高度に安定して  
いる。[信頼性を示唆]
- 大学生間の能力差を適切に検知している、と考  
えられる。[妥当性を示唆]
- 語彙、リスニング、リーディングに関わる3因子  
構造でうまく説明でき、問題形式に合った得点構  
造である。[妥当性を示唆]



# 参考：VELCテストの基準関連妥当性 トライアル時データに基づくもの

- TOEICリスニング、リーディングを目標変数、VELCのL1, L2, L3, R1, R2, R3を予測変数としたステップワイズ重回帰分析 (N = 375)



# TOEICスコアの予測

	TOEIC L	TOEIC R	TOEIC Total
予測変数	L1, L2, L3, R3	L1, L2, L3, R1, R2, R3	左の2つの 単純合計
決定係数	58%	64%	68%
重相関	0.76	0.80	0.82
標準残差	50	51	85

特段、TOEICの予測を目的として開発していない70分、120問のテストで、TOEICの分散の7割近くを説明することが判明した。

# 予測のための回帰式

## TOEICリスニング

$$\begin{aligned} \text{TOEIC listening} = & \\ & -74.88592 + \\ & 0.0753525 * L1 + \\ & 0.1986157 * L2 + \\ & 0.2481834 * L3 + \\ & 0.1187556 * R3 \end{aligned}$$

## TOEICリーディング

$$\begin{aligned} \text{TOEIC reading} = & \\ & -199.5994 + \\ & 0.0746046 * L1 + \\ & 0.0792655 * L2 + \\ & 0.1483208 * L3 + \\ & 0.1086506 * R1 + \\ & 0.1735139 * R2 + \\ & 0.2118343 * R3 \end{aligned}$$