2021.10.31

中国心理学会 Chinese Psychological Society 北京師範大學 BEIJING NORMAL UNIVERSITY 中国基础教育质量监测协同创新中心

Stage Adaptive and Timed Item Selection Methods for Computerized Classification Test

Yingshi Huang [1] He Ren[1] Ping Chen[1]

[1]Collaborative Innovation Center of Assessment for Basic Education Quality, Beijing Normal University

The problem of **exposure control** faced by CCT

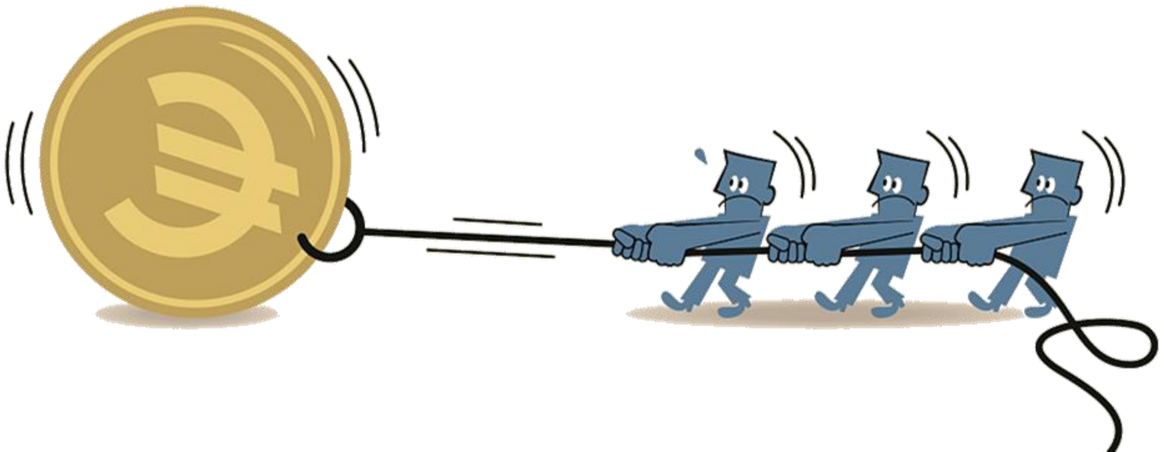
- Computerized classification testing (CCT):

✓divide students into different groups

- The traditional item selection method:

✓maximize Fisher information at the cut score (MFC)

➡ all examinees will be presented same items



The decision-making requirement varies across abilities

- in Luo et al., 2017 APM

- the “near-cut” examinees require more information
- while the “away-cut” examinees require less

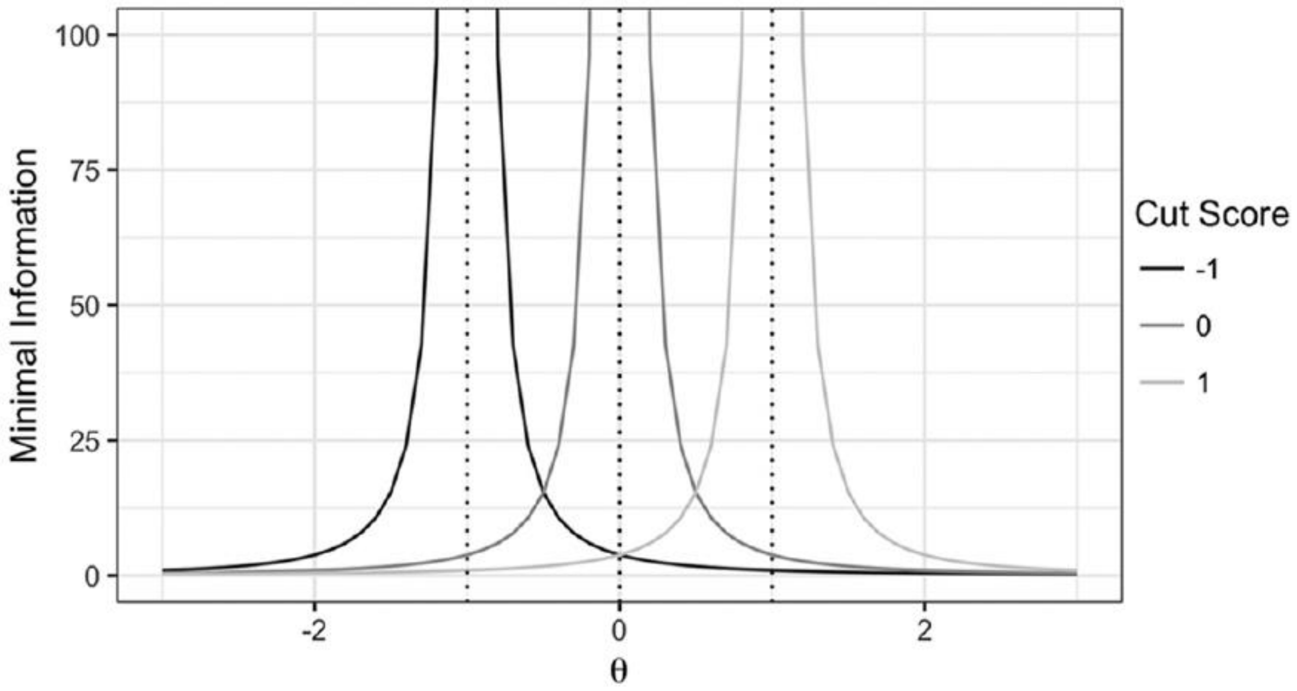


Figure 1. The minimum information requirements under the 95% confidence interval (α = .05) stopping rule.

Another increasing concern is **time control**

- The traditional MFC:

✓focus on item information but pays little attention to the response time

✓different examinees will finish the test at different time points as long as they satisfy the specific stopping rule

➡ long and uneven test-taking time

Research objectives

1. propose the **stage adaptive item selection design (SAI)** that makes the current need for decision making compatible with the percentile rank of item information;

2. optimize the “MFC per unit of time” framework (modified timed-MFC) and **put forward the timed**-SAI **method**;

3. expand these newly-proposed methods to **multidimensional scenario**.

Response and Response Time Models

- **Response Model**

✓three-parameter logistic item response model (IRT):

$$P_j(\theta) = Pr\left(X_j = 1|\theta\right)$$
$$= c_j + \left(1 - c_j\right)/\left[1 + exp\left(-a_j\left(\theta - b_j\right)\right)\right]$$

Fisher information →

$$FI_j(\theta) = \frac{\left[P_j'(\theta)\right]^2}{P_j(\theta)\left[1 - P_j(\theta)\right]}$$

✓compensatory multidimensional IRT:

$$P_j(\boldsymbol{\theta}) = Pr\left(X_j = 1|\boldsymbol{\theta}\right)$$
$$= c_j + \left(1 - c_j\right)/\left[1 + exp\left(-\left(\boldsymbol{a}_j^T\boldsymbol{\theta} + d_j\right)\right)\right]$$

Fisher Information (matrix) →

$$FI_j(\boldsymbol{\theta}) = \frac{\left[1 - P_j(\boldsymbol{\theta})\right]\left[P_j(\boldsymbol{\theta}) - c_j\right]^2}{P_j(\boldsymbol{\theta})\left[1 - c_j\right]^2}\boldsymbol{a}_j\boldsymbol{a}_j^T$$

with $\boldsymbol{\theta} = (\theta_1, \theta_2, \dots, \theta_m)$

Response and Response Time Models

- **Response Time Model**

✓log normal model:

$$f\left(t_{ij}\middle|\tau_i\right) = \frac{1}{t_{ij}\sqrt{2\pi(1/\alpha_j)^2}} e^{-[\log t_{ij}-(\beta_j-\tau_i)]^2/[2(1/\alpha_j)^2]}$$

Expected response time →

$$E\left(T_{ij}\middle|\tau_i\right) = e^{\beta_j-\tau_i+1/(2{\alpha_j}^2)}$$

✓hierarchical framework model (joint multivariate normal distribution):

Person parameter → $\boldsymbol{\psi}_P \sim MVN(\boldsymbol{\mu}_P, \boldsymbol{\Sigma}_P)$

Item parameter → $\boldsymbol{\psi}_\Gamma \sim MVN(\boldsymbol{\mu}_\Gamma, \boldsymbol{\Sigma}_\Gamma)$

Item Selection Designs in CCT

- **Traditional Item Selection Method**

✓unidimensional scenario:

→ maximize Fisher information at the cut score

$$i_k = \underset{j \in R_{k-1}}{\operatorname{argmax}} FI_j(\theta_c)$$

✓multidimensional scenario:

→ maximize the determinant of the Fisher information matrix at the cut score on **the reference composite (RC)**

Item Selection Designs in CCT

- the Reference Composite Method

✓the main goal:

→ project the multidimensional $\boldsymbol{\theta}_s$ from the ability space onto the line with a specific **direction** that best measured at the test level

⬇

the eigenvector of $\boldsymbol{a}_j^T \boldsymbol{a}_j$ that corresponds to the largest eigenvalue

direction cosines:

the lth element ($l = 1, \ldots, m$) of the eigenvector represent the cosine of the angle between the RC and the dimension axes ($\alpha_{\xi l}$)

Item Selection Designs in CCT

- the Reference Composite Method

✓transform the ξ_c to the multidimensional space:

$$\boldsymbol{\theta}_c = \xi_c cos\boldsymbol{\alpha}_\xi$$

✓maximize the Fisher information determinant at the cut scores (similar to the D-optimal):

$$i_k = \underset{j \in R_{k-1}}{\text{argmax}} \left\{ det\left(\sum_{i=1}^{k-1} FI_{k-1}(\boldsymbol{\theta}_c) + FI_j(\boldsymbol{\theta}_c) \right) \right\}$$

Item Selection Designs in CCT

- **New Item Selection Methods**

1. the Stage Adaptive Item Selection Method

✓define the need at current stage:

$$s = \left|\frac{\hat{\theta} - \theta_c}{range\ of\ \hat{\theta}}\right| = \frac{1}{(-4-4) = 8} \times \left|\hat{\theta} - \theta_c\right|$$

✓the stage adaptive index:

$$SAI = \exp\{-|r[FI_j(\theta_c)] + w \times s - 1|\}$$

$$i_k = \underset{j \in R_{k-1}}{\operatorname{argmax}} SAI$$

– the exponential form makes it convenient to combine with other constraints (response time limitation or content balance etc.)

– *w*: the weighting parameter

Item Selection Designs in CCT

1. the Stage Adaptive Item Selection Method

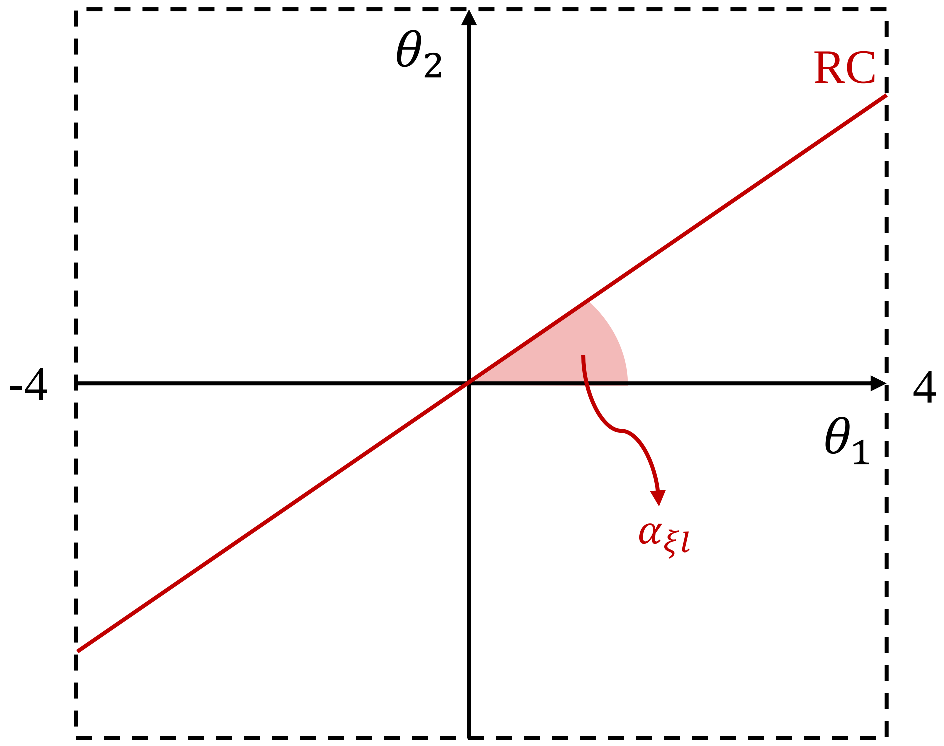
✓under the multidimensional scenario:

$$s_{RC} = \left| \frac{\hat{\xi} - \xi_c}{range\ of\ RC} \right|$$

$$\frac{4}{\frac{1}{2} \times range_{RC}} = max(cos\boldsymbol{\alpha}_{\xi})$$

$$SAI = \exp\{-|r[det(\boldsymbol{\theta}_c)_{RC}] + w \times s_{RC} - 1|\}$$

$$i_k = \underset{j \in R_{k-1}}{\operatorname{argmax}} SAI$$



Item Selection Designs in CCT

2. the Item Selection Methods with Response Time

✓Existing method: timed-MFC by Sie et al. (2015)

$$i_k = \underset{j \in R_{k-1}}{\operatorname{argmax}} \frac{FI_j(\theta_c)}{E\left(T_j \middle| \hat{\tau}_{k-1}\right)}$$

✓**Modified timed-MFC:** combine the idea proposed by Choe et al. (2018)

$$i_k = \underset{j \in R_{k-1}}{\operatorname{argmax}} \frac{FI_j(\theta_c)}{\left|E\left(T_j \middle| \hat{\tau}_{k-1}\right) - v\right|}$$

Consider the impact of speed →

$$\beta_j + \frac{1}{2{\alpha_j}^2} = \hat{\tau}_{k-1} + lnv$$

✓Put forward the **timed-SAI method**:

$$i_k = \underset{j \in R_{k-1}}{\operatorname{argmax}} \frac{SAI}{\left|E\left(T_j \middle| \hat{\tau}_{k-1}\right) - v\right|}$$

Simulation studies

- **Study 1**

✓To display the nature of the SAI design and investigate its performance

- **Study 2**

✓To ascertain the potential of the two new timed designs (i.e., the modified timed-MFC and the timed-SAI designs) on test time controlling

- **Study 3**

✓To investigate the performance of the multidimensional extended version of the new designs

Simulation studies

- **Study 1**

✓item parameter:

$a_j \sim U(1.0,2.5)$

$b_j \sim N(0,1)$

$c_j \sim beta\,(2,10)$

✓item bank scale: 500 items

✓test length: 10, 20 items
cut score = 0

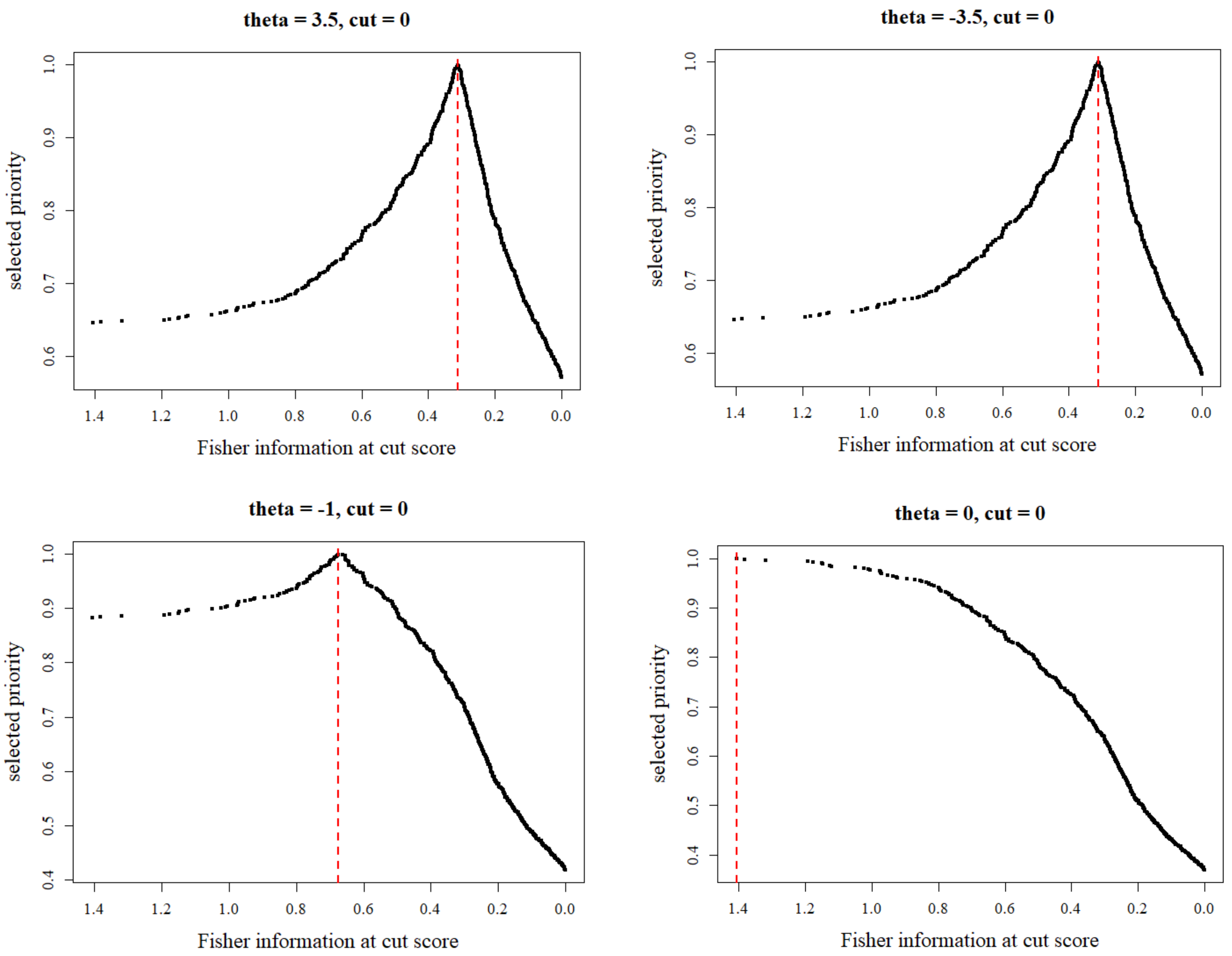
✓ability level: 29 points from -3.5 to 3.5 by 0.25 (each contains 100 examinees, 2900 in total)
ability estimation: MLE combines with EAP

✓item selection method: MFC vs SAI
weighting parameter for SAI: 0.50, 0.75, 1.00, 1.25
select the first item randomly

✓evaluation criteria (over 100 repetitions):
item exposure rate
percentage of correct classifications (PCC)

Simulation studies

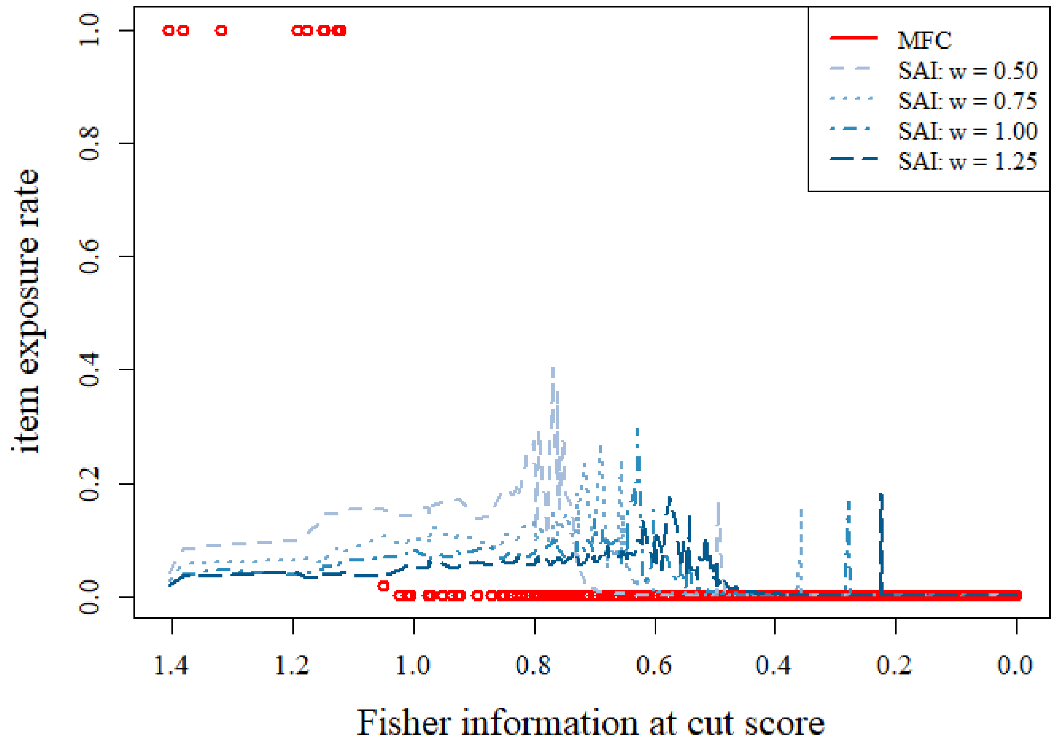
- **Study 1**



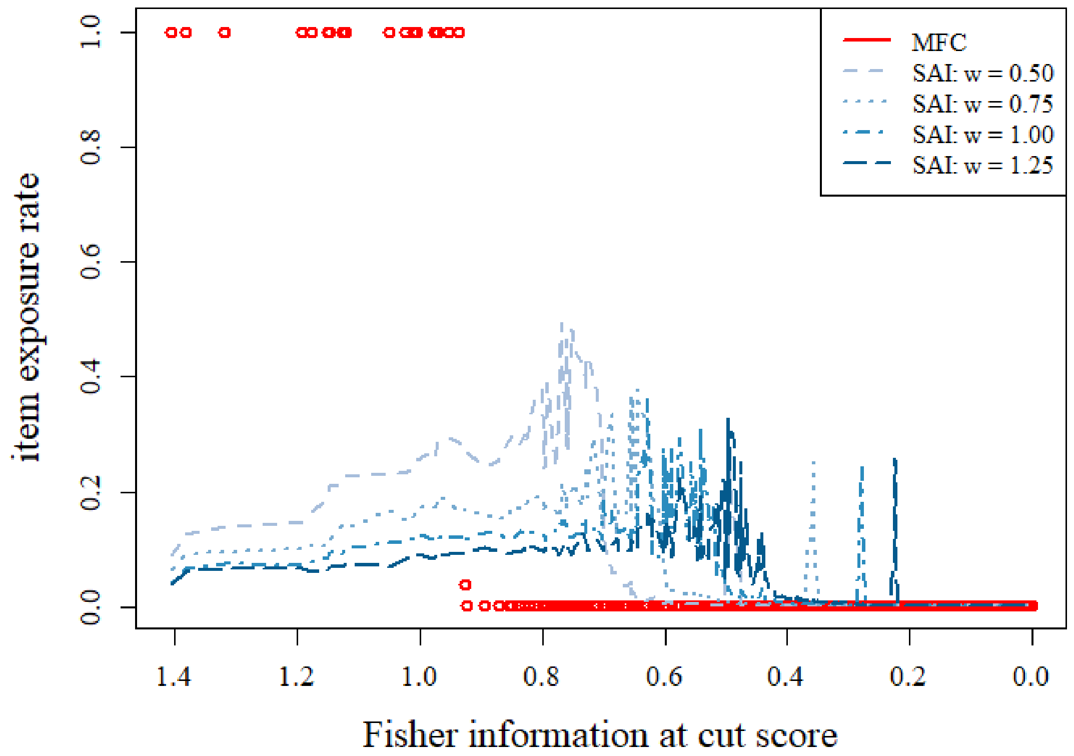
Simulation studies

- **Study 1**

✓10 items:

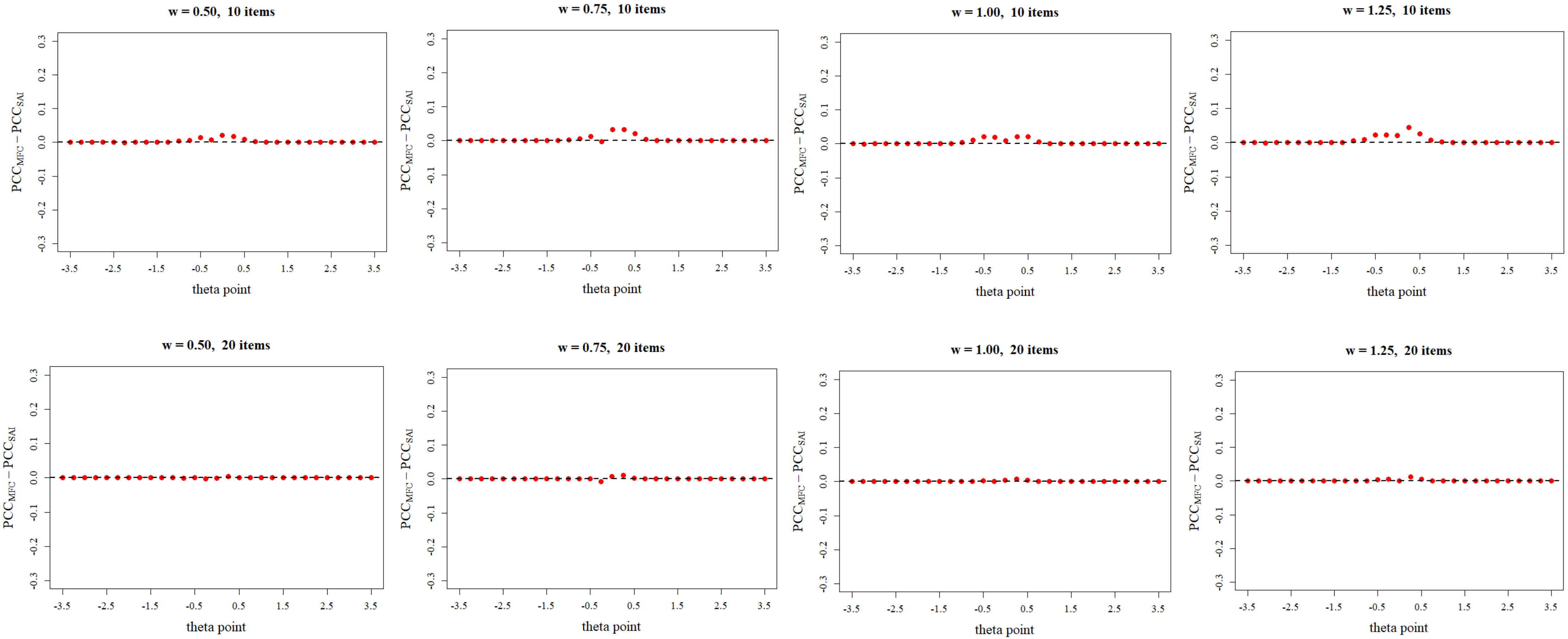


✓20 items:



Simulation studies

- **Study 1**



Simulation studies

- **Study 2**

✓item parameter:

$\alpha_j \sim U[1,3]$

$MVN(\boldsymbol{\mu}_\Gamma, \boldsymbol{\Sigma}_\Gamma)$

$\boldsymbol{\mu}_\Gamma = (\mu_b, \mu_\beta) = (0,4)$

$\boldsymbol{\Sigma}_\Gamma = \begin{bmatrix} 1 & 0.25 \\ 0.25 & 0.25 \end{bmatrix}$

✓person parameter (1000 examinees):

$MVN(\boldsymbol{\mu}_\mathrm{P}, \boldsymbol{\Sigma}_\mathrm{P})$

$\boldsymbol{\mu}_\mathrm{P} = (\mu_\theta, \mu_\tau) = (0,0)$

$\boldsymbol{\Sigma}_\mathrm{P} = \begin{bmatrix} 1 & 0.50 \\ 0.50 & 1 \end{bmatrix}$

✓item selection method:

MFC vs modified timed-MFC vs timed-SAI

weighting parameter for SAI: $w = 1$

center parameter for timed method:

v = seq(0, $e^{\mu_\beta - \mu_\tau}$, length = 30) + 5 additional points

(select the first item randomly)

✓evaluation criteria (over 100 repetitions):

test-taking time

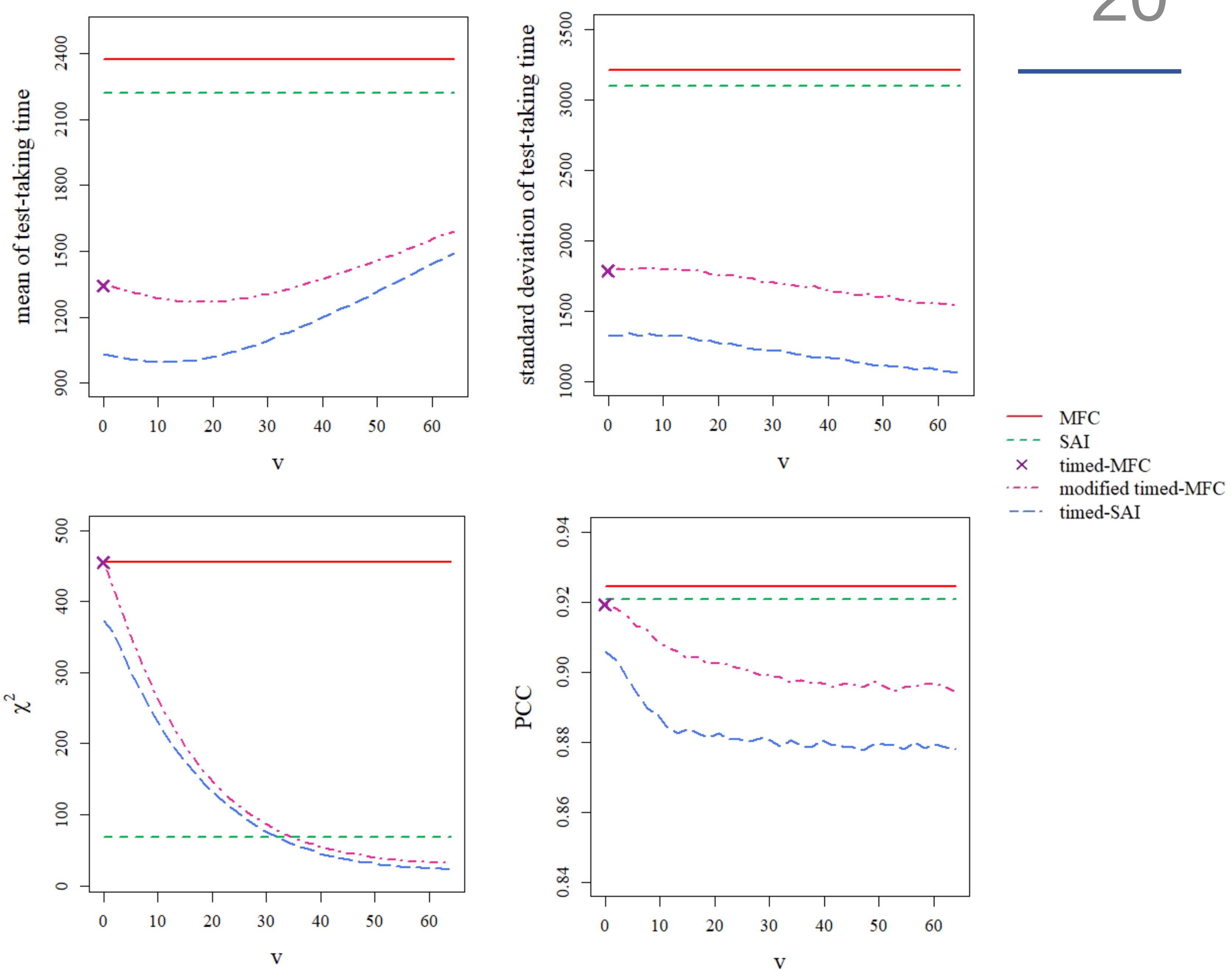
percentage of correct classifications (PCC)

the χ^2 statistic

$$\chi^2 = \sum_{j=1}^{J} \frac{\left(er_j - \overline{er_j}\right)^2}{\overline{er_j}}$$

Simulation studies

- **Study 2**



Simulation studies

- **Study 3**

✓item parameter:

$MVN(\boldsymbol{\mu}_\Gamma, \boldsymbol{\Sigma}_\Gamma)$

$\boldsymbol{\mu}_\Gamma = (\mu_d, \mu_\beta) = (0,4)$

$\boldsymbol{\Sigma}_\Gamma = \begin{bmatrix} 1 & 0 \\ 0 & 0.25 \end{bmatrix}$

$\log(a_{jl}) \sim N(0, 1)$

with $a_{jl} \in (0.2, 2.5)$

✓person parameter (1000 examinees):

$MVN(\boldsymbol{\mu}_\mathrm{P}, \boldsymbol{\Sigma}_\mathrm{P})$

$\boldsymbol{\mu}_\mathrm{P} = (\mu_{\theta_1}, \mu_{\theta_2}, \mu_{\theta_3}, \mu_\tau) = (0,0,0,0)$

$\boldsymbol{\Sigma}_\mathrm{P} = \begin{bmatrix} 1 & 0.3 & 0.3 & 0.5 \\ 0.3 & 1 & 0.3 & 0.5 \\ 0.3 & 0.3 & 1 & 0.5 \\ 0.5 & 0.5 & 0.5 & 1 \end{bmatrix}$

✓test length: 30 items

cut score: $\xi_c = 0$, thus, $\boldsymbol{\theta}_c = (0,0,0)$

✓item selection method:

RC vs SAI vs timed-RC vs timed-SAI

weighting parameter for SAI: $w = 1$

center parameter for timed method:

$v = \{1, e^{\mu_\beta - \mu_\tau}/2\}$

(select the first four items randomly)

estimation (constrained MLE): $\widehat{\boldsymbol{\theta}}_{MLE,i} = \underset{\boldsymbol{\theta} \in [-4,4] \times [-4,4] \times [-4,4]}{\operatorname{argmax}} \{log[L(\boldsymbol{\theta} | \boldsymbol{y}_i)]\}$

Simulation studies

- **Study 3**

		PCC	χ^2 statistic	mean of test time	SD of test time
$v = 1$	RC	0.909	396.924	3754.538	4581.262
	SAI	0.901	23.979	3464.855	4227.865
	timed-RC	0.892	390.863	1552.895	1907.627
	timed-SAI	0.897	332.325	1621.848	1992.096
$v = \frac{e^{\mu_\beta - \mu_\tau}}{2}$	timed-RC	0.887	100.897	1719.759	1806.634
	timed-SAI	0.892	93.167	1760.304	1892.445

Discussion

The present research **enhances the item selection method** for both unidimensional and multidimensional scenarios in two points:

1. assigning items adaptive to the need at the current stage

2. restricting the expected RT across all examinees to the same level

Thanks for listening!

黄颖诗 h_yingshi@163.com

Item Selection Designs in CCT

1. the Stage Adaptive Item Selection Method

✓there are serval versions for the definition of *s*:

$$s_1 = \left|\frac{\hat{\theta} - \theta_c}{range\ of\ \hat{\theta}}\right| = \frac{1}{8} \times \left|\hat{\theta} - \theta_c\right|$$

faster, but rough

$$s_2 = \left|\frac{\hat{\theta}-\theta_c}{4-\theta_c}\right| \text{ for } \hat{\theta} > \theta_c \text{ and } s_2 = \left|\frac{\hat{\theta}-\theta_c}{-4-\theta_c}\right| \text{ for } \hat{\theta} < \theta_c$$

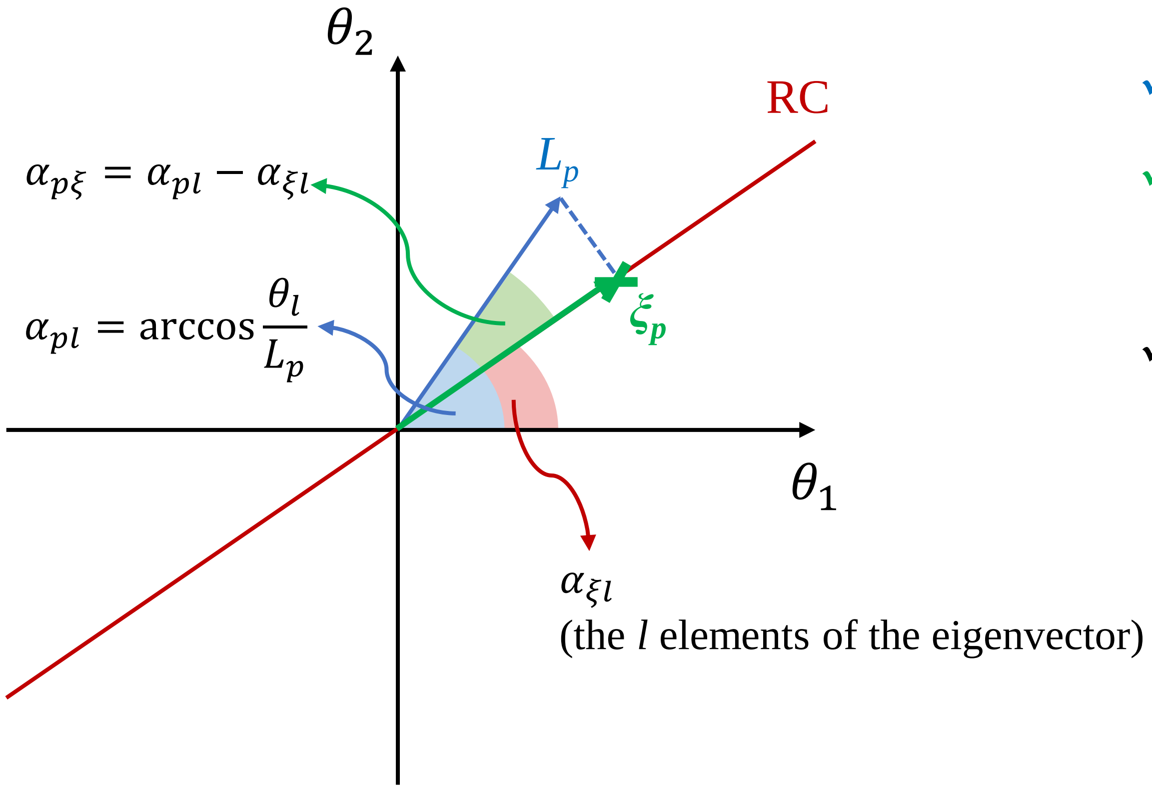
$$s_3 = \left|\frac{\hat{\theta} - \theta_c}{Z_\alpha \times SEM}\right| = \left|\frac{\hat{\theta} - \theta_c}{1.96 \times \frac{1}{\sqrt{\sum_{j=1}^{k} FI_j(\hat{\theta})}}}\right|$$

precise, but with larger calculation

- it is plausible that three designs can shift from each other by "*w* × *s*" (*w* > 1)

Item Selection Designs in CCT

- Take a two dimensional scenario as example



✓ L_p : the length of ability vector for the examinee p, $L_p = \|\boldsymbol{\theta}\|$

✓ ξ_p : the projection of abilities on the RC, $\xi_p = L_p cos\alpha_{p\xi}$

✓ Let ξ_c be the cut-off point on the RC, thus:

pass the exam if $\xi_p \geq \xi_c$; fail otherwise

Future directions

1. combine the current design with other methods to **meet various constraints simultaneously**

2. generalize to the variable-length (VL) test setting as well as the multiple categories scenario

3. explore the effect of speed on **decision precision**