American Psychological Association
Psychological Review

Cognitive Psychology Meets Psychometric Theory: On the Relation Between Process Models for Decision Making and Latent Variable Models for Individual Differences

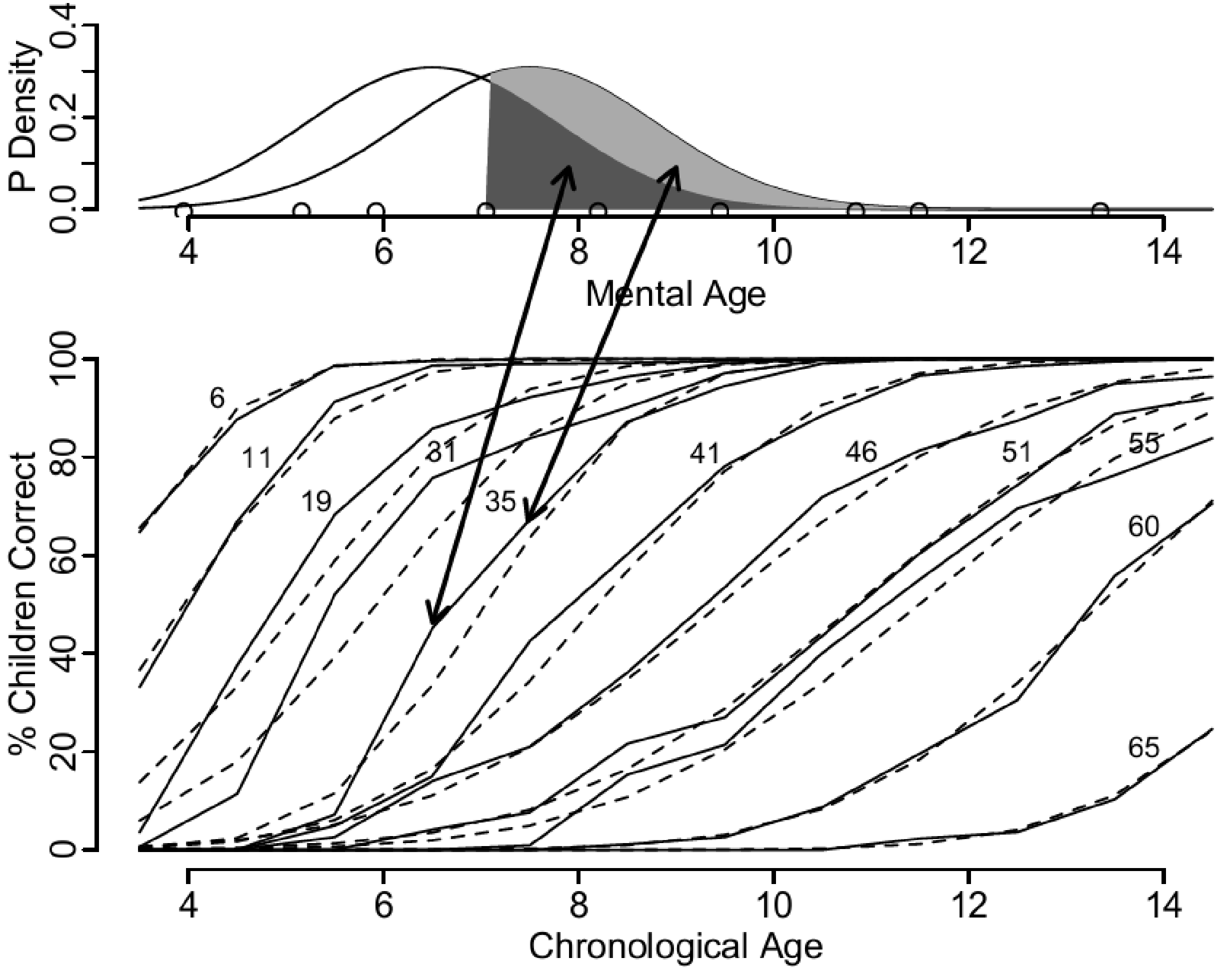
Han L. J. van der Maas and Dylan Molenaar
University of Amsterdam

Gunter Maris
Cito, Arnhem, The Netherlands, and University of Amsterdam

Rogier A. Kievit and Denny Borsboom
University of Amsterdam

Reporter: Huang Yingshi

Let’s begin with the item response theory



Thissen & Steinberg *C/EJEME* (2020)

IRT has proven to be extremely useful in test analysis

- the item response model (e.g., 2PLM)

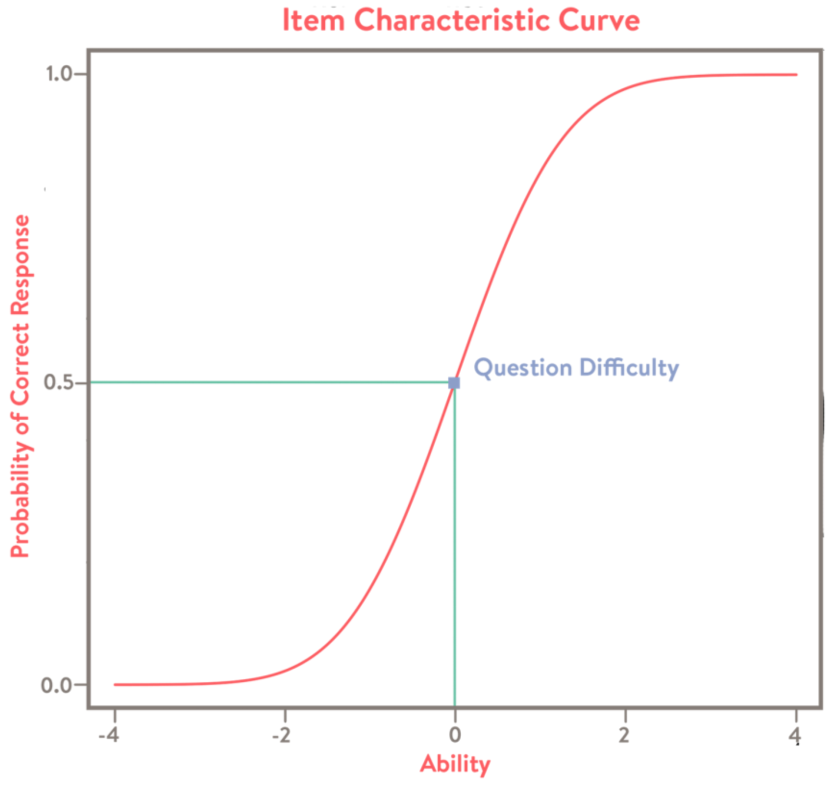
$$P_+ = \frac{e^{\alpha(\theta-\beta)}}{1 + e^{\alpha(\theta-\beta)}}$$

✓ equating

✓ computer adaptive testing

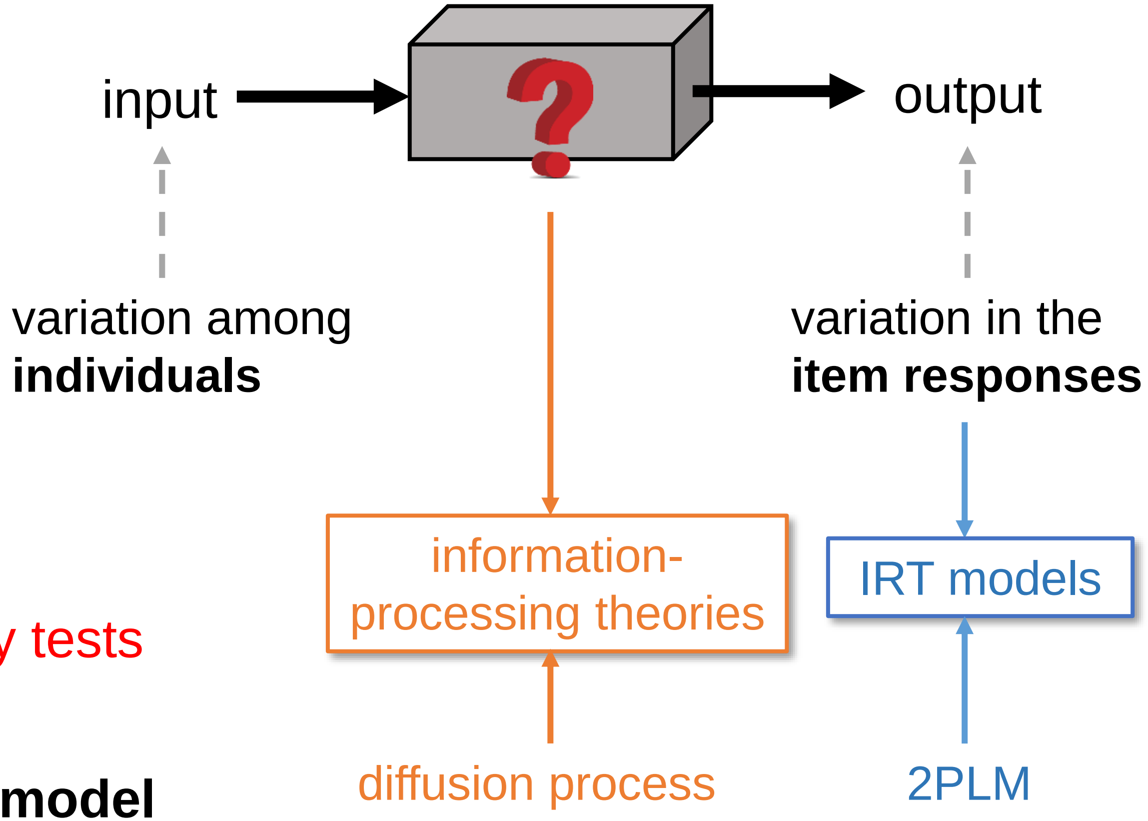
✓ the investigation of differential item functioning

✓ …



But how these item responses are generated?

$$P_{+} = \frac{e^{\alpha(\theta-\beta)}}{1 + e^{\alpha(\theta-\beta)}}$$



1. Tuerlinckx & De Boeck (2005)

   attitude and personality tests vs. ability tests

   - **a new IRT model: the Q-diffusion model**

The decision making process

- Which one would you choose for dinner?

Option A

Option B

Diffusion model: sequential sampling of information

- Item response processes require the respondent to **make a decision**

In the two-choice response task:
collect evidence for the response options

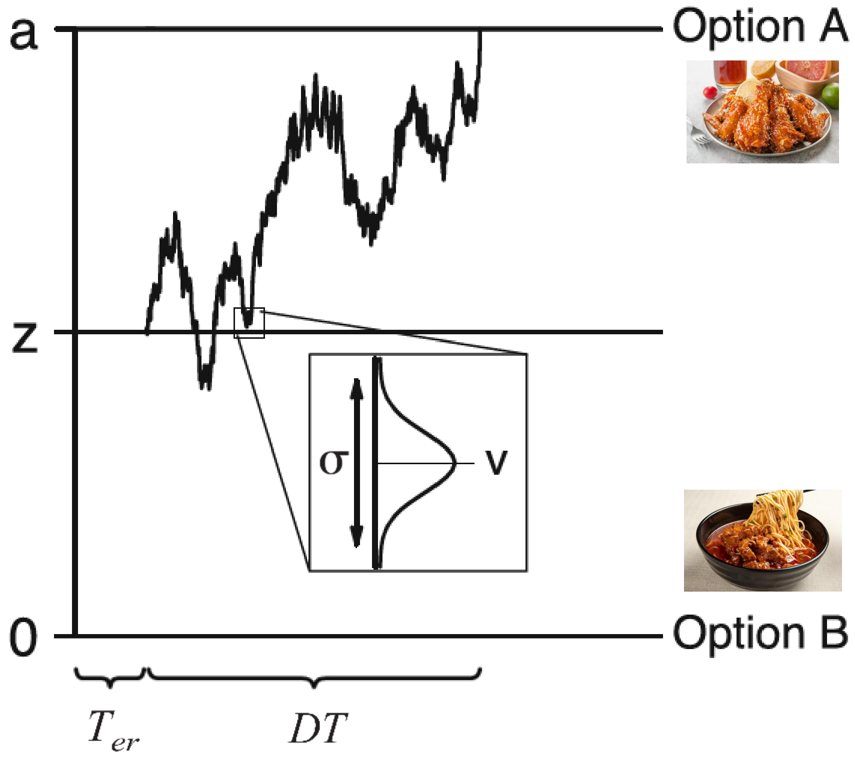
✓v: drift rate

✓σ: diffusion coefficient

✓a: boundary separation

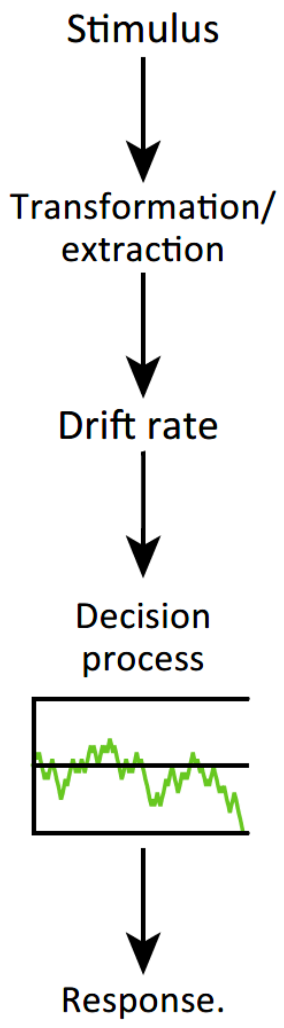
✓z: starting point

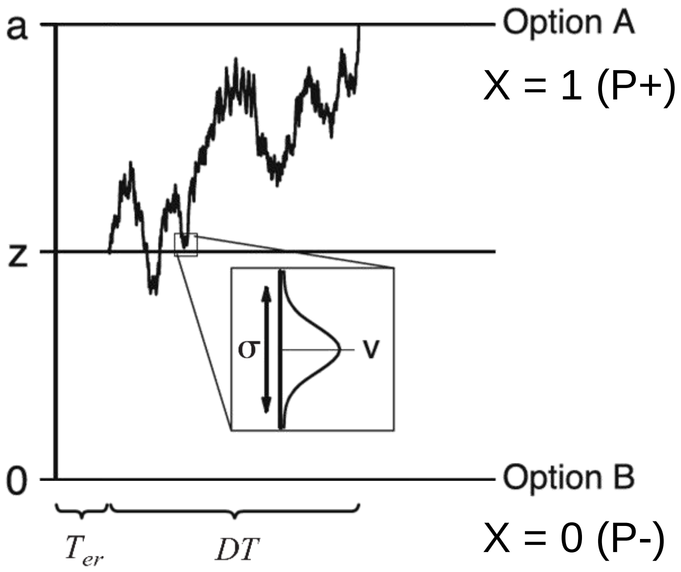
Response time T = nondecision time T_{er} + decision time DT



Boehm et.al. *PSYCHOMETRIKA* (2021)

To model the probability of a correct response





- $$f_{X,T}(x,t) = \frac{\pi\sigma^2}{a^2} \exp\left(\frac{(ax - z)v}{\sigma^2} - \frac{v^2}{2\sigma^2}(t - T_{er})\right) \times \sum_{m=1}^{\infty} m \sin\left(\frac{\pi m(ax - 2zx + z)}{a}\right) \times \exp\left(-\frac{1}{2}\frac{\pi^2\sigma^2 m^2}{a^2}(t - T_{er})\right)$$
- $$P_+ = P(X = 1) = \frac{e^{-2zv} - 1}{e^{-2av} - 1} = \frac{e^{av}}{1 + e^{av}}$$ in an unbiased decision process (i.e., z = a/2)

Tuerlinckx & De Boeck (2005): $a = \alpha$

$v = \theta - \beta$

person part (θ), item part (β)

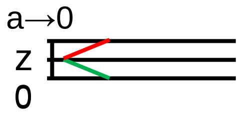
Implications of the diffusion interpretation of IRT models

- response process

✓ the probability distribution of **item responses** + the distribution of **response times**

- properties of the response times

1. time limit reduces → boundary separation approaches 0 → P+ = 0.5 for all θ_s



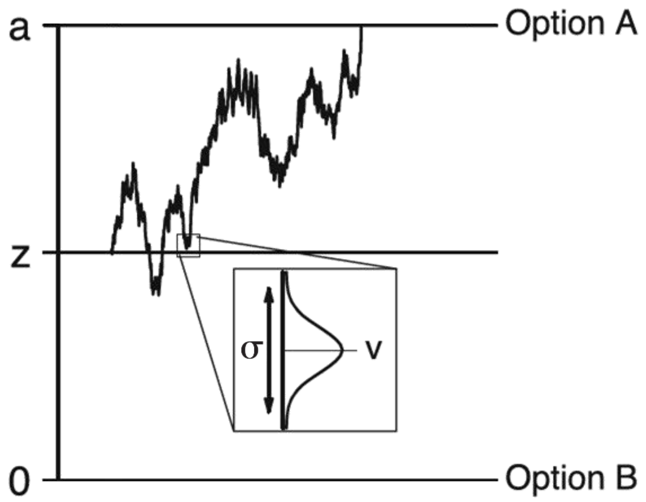
Ability Testing:

- the dichotomies in the data result from **scoring (incorrect–correct)** rather than from a **two-choice situation**
- for M response options: 1/M

Implications of the diffusion interpretation of IRT models

- properties of the response times

  2. the effect of changes in v



- In diffusion model: slowest when v ≈ 0
- In IRT model: $v = \theta - \beta$

  when θ = β → slowest
  when θ >> β & θ << β → vary fast

Personality and attitude items

	agree	*disagree*
the death penalty is allowed	☐	☐

Ability tests

low ability individuals give the incorrect response
as fast as high ability individuals give the correct response

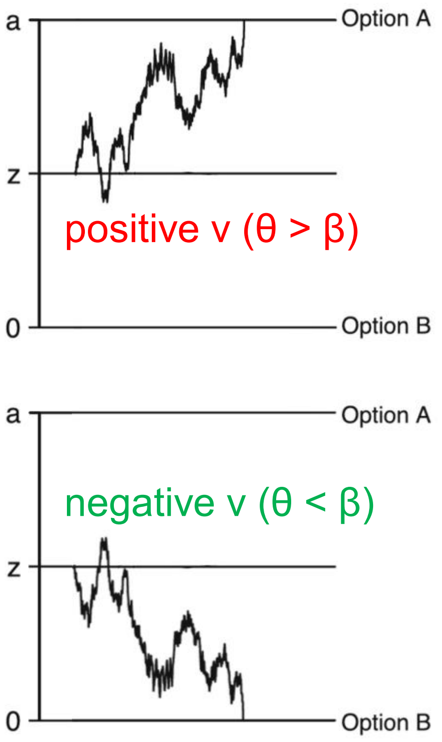
Implications of the diffusion interpretation of IRT models

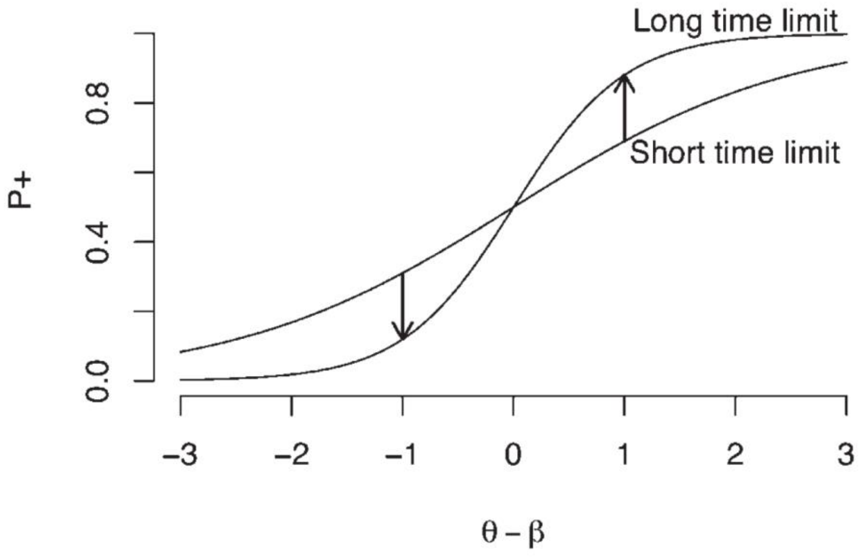
- properties of the response times

  3. item discrimination (which determined by boundary separation)

$a \propto P+$

$a \propto P-$





Personality and attitude tests:

- ✓ thinks about his/her position **longer** → select the option that **best fits** his/her latent state

Ability tests:

- ✓ increasing the time limit will **increase P+**
- ✓ the worst they can do is **guess**

$$P_{+}^{3PL} = c_j + (1 - c_j)\, P_{+}^{2PL}$$

a simpler and more fundamental solution is possible

Consider the deep structure of the ability concept

- In current psychometric theory
  - individual differences:
    the relation between the examinee *i* and other test takers

- In the diffusion model
  - Parameters at play in the actual process that a single individual follows when answering a test item

What is ability at the level of an individual?

Three characteristics for simple abilities and the tasks

1. the ability: present or absent (cannot be negative)
   - a capacity to do something

2. the difficulties: essentially positive as well
   - any task that can be said to measure this ability requires some of the ability

3. the drift rate: always positive
   - task can be carried out by any individual who possesses the ability if only the individual is given sufficient time
   - i.e., P+ = 1 if time limits are absent

Introduce individual differences into the model

1. if there are time limits:
   - high ability examinee: higher probability for success

2. if there are no time limits:
   - high ability examinee: complete that task faster

3. in the general situation:
   - speed–accuracy trade-off:
     individual differences in both the probability and the time

What characteristics should the new model possess?

- the diffusion model
  - no **clear separation** between person and item parameters

For drift rate:

$$v = f(v^p, v^i)$$ ➡ ability & difficulty

For boundary separation:

$$a = g(a^p, a^i)$$ ➡ response caution & time pressure

1. v and a must be positive
2. P+ monotonically increasing in v^p and monotonically decreasing in v^i
3. P+ = 1 when: v^p approaches infinity or v^i approaches 0
4. P+ = chance level when: v^i approaches infinity or v^p approaches 0

The positive ability model

- the sequential sampling based item response model
  - Newtonian relation:
    v = P/F
    speed (drift rate) = power (ability) / force (difficulty)

$$v = v^p / v^i \qquad a = a^p / a^i \quad \Rightarrow \quad P_+ = \frac{e^{av}}{1 + e^{av}} = \frac{e^{\frac{a_k^p v_k^p}{a_j^i v_j^i}}}{1 + e^{\frac{a_k^p v_k^p}{a_j^i v_j^i}}}$$

✓ positive v & positive a
✓ inverse proportion
✓ time limit:
larger (P+ → chance level 0.5) & smaller (P+ → 1)

The positive ability model

- tests with multiple response options
  - nominal response model

$$P_m = \frac{e^{\beta^*_m + \alpha^*_m \theta}}{\sum_{k=1}^{M} e^{\beta^*_k + \alpha^*_k \theta}}$$

Which one would you choose for dinner? (M = 3)


Option A


Option B


Option C

assume:

1. m = M is the correct answer
2. incorrect alternatives are all equally attractive (set $\alpha^*_1 \ldots \alpha^*_{M-1}$ and $\beta^*_1 \ldots \beta^*_{M-1}$ to zero)

$$P_m = \frac{e^{\beta^*_m + \alpha^*_m \theta}}{(M-1)e^{0+0\theta} + e^{\beta^*_m + \alpha^*_m \theta}} = \frac{e^{\beta^*_m + a^*_m \theta}}{(M-1) + e^{\beta^*_m + \alpha^*_m \theta}} = \frac{\frac{e^{\beta^*_m + \alpha^*_m \theta}}{e^{\ln(M-1)}}}{\frac{e^{\ln(M-1)}}{e^{\ln(M-1)}} + \frac{e^{\beta^*_m + \alpha^*_m \theta}}{e^{\ln(M-1)}}} = \frac{e^{\beta^*_m + \alpha^*_m \theta - \ln(M-1)}}{1 + e^{\beta^*_m + \alpha^*_m \theta - \ln(M-1)}}$$

The positive ability model

- tests with multiple response options
  - apply to the positive ability model

$$P_m = \frac{e^{\beta_m^* + \alpha_m^* \theta - \ln(M-1)}}{1 + e^{\beta_m^* + \alpha_m^* \theta - \ln(M-1)}}$$

rewrite as a modified 2PLM (set $\alpha^* = \alpha$ and $\beta^* = -\beta\alpha$)

$$P_+ = \frac{e^{\alpha(\theta-\beta) - \ln(M-1)}}{1 + e^{\alpha(\theta-\beta) - \ln(M-1)}}$$

$$\theta_k = a_k^p v_k^p$$
$$\alpha_j^* = \alpha_j = 1/a_j^i v_j^i$$
$$\beta^* = -\ln(M_j - 1)$$

$$P_+ = \frac{e^{\frac{a_k^p v_k^p}{a_j^i v_j^i} - \ln(M_j - 1)}}{1 + e^{\frac{a_k^p v_k^p}{a_j^i v_j^i} - \ln(M_j - 1)}}$$

$$\theta_k = a_k^p v_k^p$$
$$\alpha_j = 1/a_j^i v_j^i$$
$$\beta_j = \ln(M_j - 1) a_j^i v_j^i$$

$$P_+ = \frac{e^{\alpha(\theta-\beta)}}{1 + e^{\alpha(\theta-\beta)}}$$

A general framework

- the Q-diffusion model (QM)
  - the quotient model on a diffusion model basis

$$P_{+} = \frac{e^{\frac{a_k^p v_k^p}{a_j^i v_j^i} - \ln(M_j - 1)}}{1 + e^{\frac{a_k^p v_k^p}{a_j^i v_j^i} - \ln(M_j - 1)}} = \frac{e^{\frac{a_k^p v_k^p}{a_j^i v_j^i}}}{M - 1 + e^{\frac{a_k^p v_k^p}{a_j^i v_j^i}}}$$

when $\theta = a^p v^p$, $\beta = a^i v^i$, and $K = M - 1$

$$P_{+} = \frac{e^{\theta_k/\beta_j}}{K + e^{\theta_k/\beta_j}},\ \theta_k \geq 0,\ \beta_j > 0$$

represent the absence of ability

Use response times to separate item and person factors

$$\theta_k = a_k^p v_k^p$$

- two factors have different effects on response time
  - response caution ∝ RT
  - information power ∝ 1/RT

- response time distribution

$$\log(RT_{kj}) \sim \text{normal}(\mu_{\text{kj}}, \sigma^2_{\text{kj}})$$

$$u_{kj} = \log[E(RT_{kj})] - \frac{1}{2}\left[1 + \frac{\text{var}(RT_{kj})}{E(RT_{kj})^2}\right]$$

$$\sigma^2_{kj} = \log\left[1 + \frac{\text{var}(RT_{kj})}{E(RT_{kj})^2}\right]$$

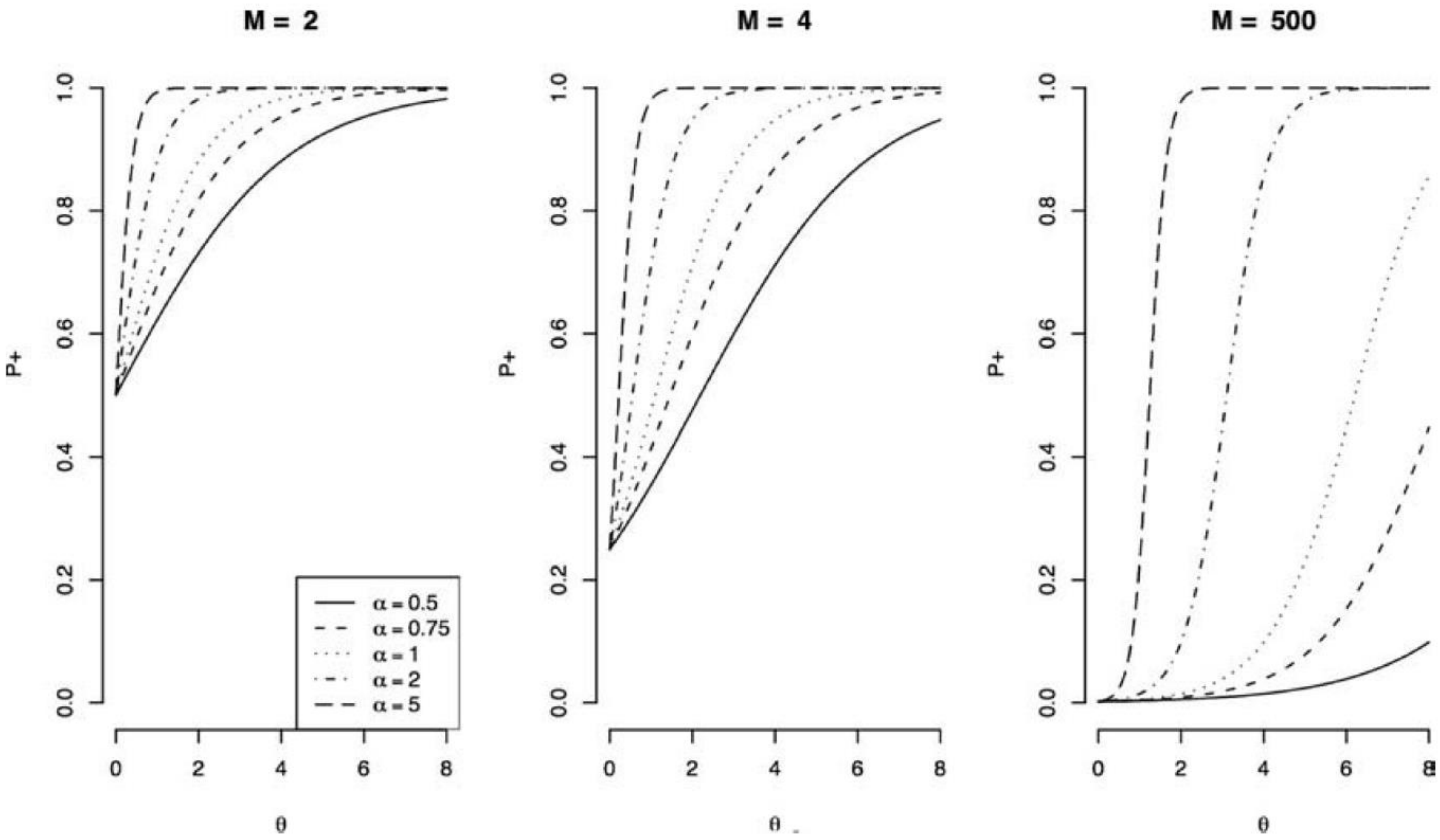
are functions of the Q-diffusion model parameters →

$$E(RT_{kj}) = \frac{a_k^p v_j^i}{2 a_j^i v_k^p} \frac{1 - e^{h_{kj}}}{1 + e^{h_{kj}}} + Ter_k$$

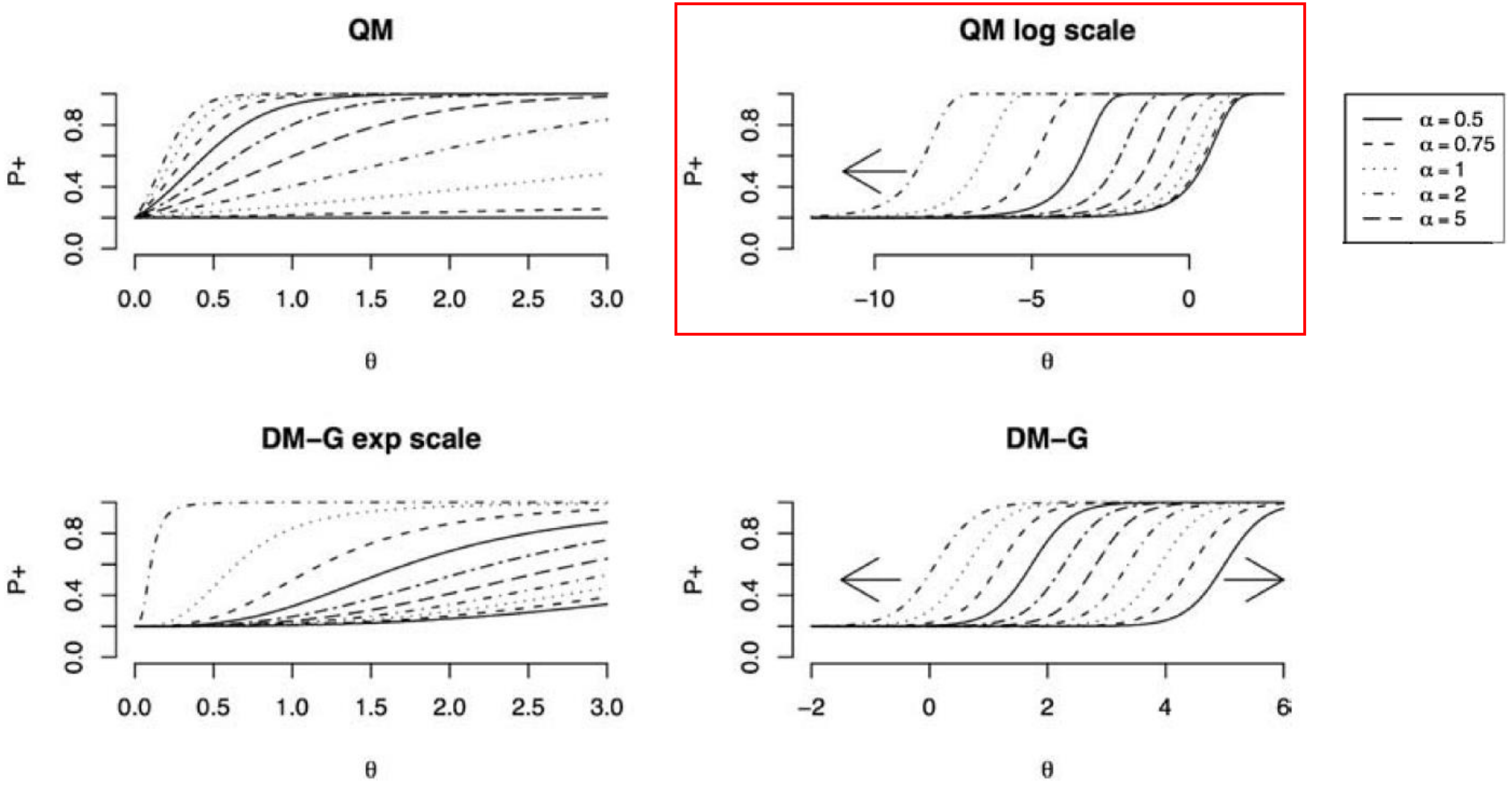
$$\text{var}(RT_{kj}) = \frac{a_k^p}{2 a_j^i}\left(\frac{v_j^i}{v_k^p}\right)^3 \left[\frac{2 h_{kj} e^{h_{kj}} - e^{2 h_{kj}} + 1}{(e^{h_{kj}} + 1)^2}\right]$$

$$h_{kj} = -\frac{v_k^p a_k^p}{v_j^i a_j^i}$$

Item characteristic curve (ICC) of the Q-diffusion model

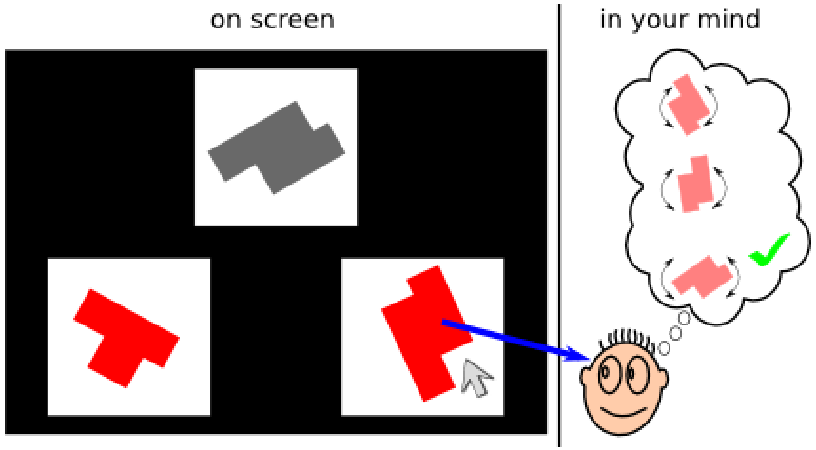


Compare with the Rasch model with guessing (DM-G)



Fit the Q-diffusion model and present two examples

- Example 1: Mental Rotation
  - 121 subjects in the context of a mental rotation task
  - Responses were dichotomous (correct vs. incorrect)
  - 10 items with three different rotation angles (50°, 100°, 150°)



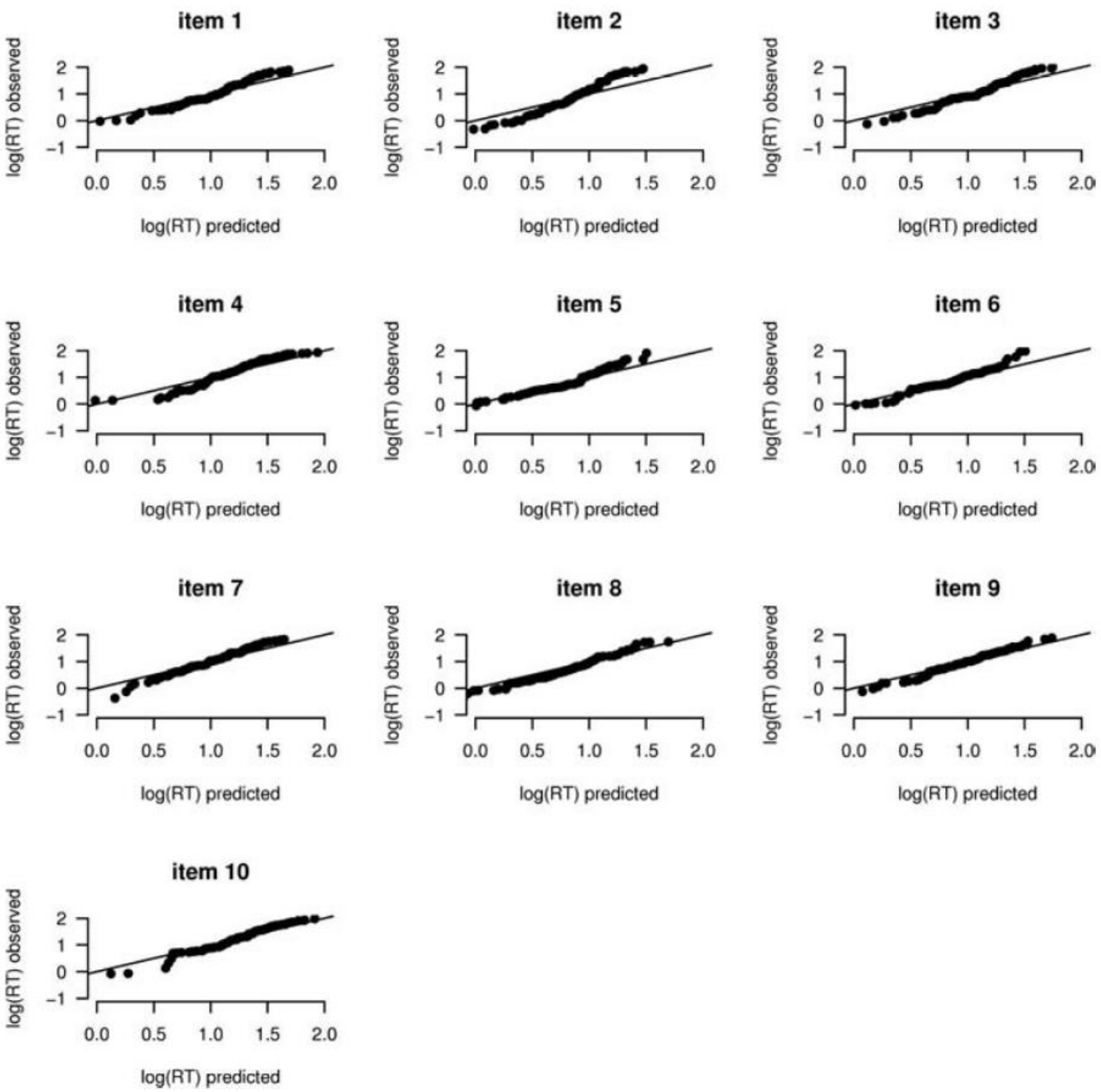
- Estimation:
  - ➢ without response time: **fit statistics**
    only for a comparison with standard IRT models
  - ➢ with response time: **predicted and observed RT**
    MCMC - sample from the posterior distribution
    (uninformative priors)

The Q-diffusion model outperforms the other IRT models

Fit Statistics for the Q-Diffusion Item Response Model and Several Standard Item Response Models

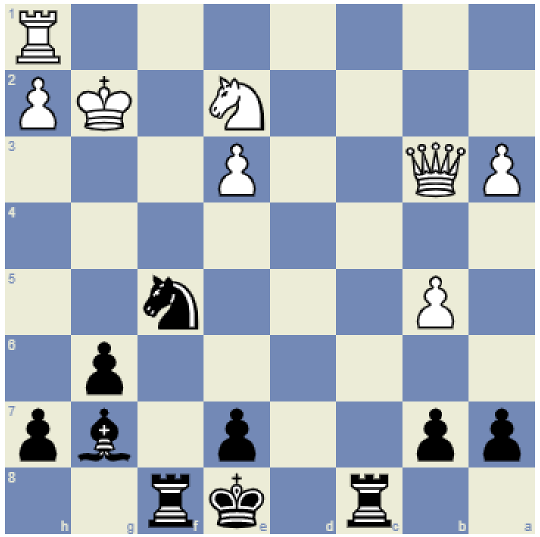
Model	−2LL	AIC	BIC
Mental rotation example			
Q-diffusion	832.2	*852.2*	*880.2*
1PL	835.1	857.1	887.9
1PL guessing	846.5	866.5	894.5
2PL	830.5	870.5	926.4
3PL full	819.2	859.2	915.1

High degree of equivalence for the mental rotation data



An open ended example with strong external criterion

- Example 2: Chess puzzles
  - a multiple-choice format with an unknown number of options
  - external criterion: Elo ratings
  - consist of many different abilities
  - 20 chess items
  - estimation: with response time (full model)



Using drift rate instead of test score improves validity

Correlations of the Standard Test Statistics, Person Estimates According to the 1PL and 2PL Models, and the Q-Diffusion Parameters Person Drift Rate (v), Response Caution (a), and Nondecision Time (T_{er}), With the Elo Ratings and Ages of Chess Players

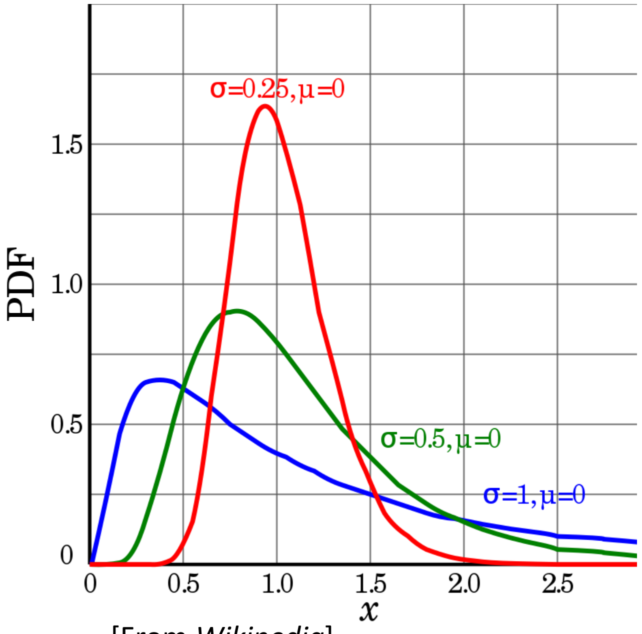
Person	Test score	Response time	1PL θ	2PL θ	v^p	a^p	T_{er}
Elo rating	0.68	−0.44	0.67	0.69	0.72	−0.38	−0.17
Age	−0.35	0.54	−0.35	−0.33	−0.34	0.24	0.60

With the accuracy data only, the 2PL model fit better

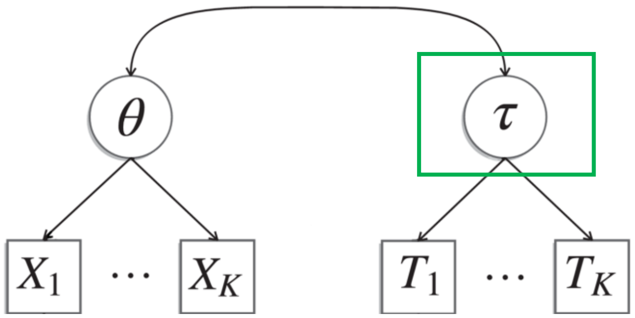
Fit Statistics for the Q-Diffusion Item Response Model and Several Standard Item Response Models

Model	−2LL	AIC	BIC
	Chess ability example		
Q-diffusion	4,263.0	4,341.0	4,479.2
1PL	4,309.3	4,351.3	4,425.7
1PL guessing	4,214.8	4,294.8	4,436.7
2PL	4,178.4	*4,258.4*	*4,400.3*
3PL full	4,164.8	4,282.8	4,491.9

Compare with the hierarchical approach



[From *Wikipedia*]



- In the hierarchical model
  - latent construct:
    response time is the ratio of amount of labor and speed

$$E[\ln(RT_{jk})] = \xi_j - \tau_k$$

- In the Q-diffusion model
  - process parameters:
    not defined by their effects on the probability of response and time

Link the Q-diffusion model to the hierarchical model

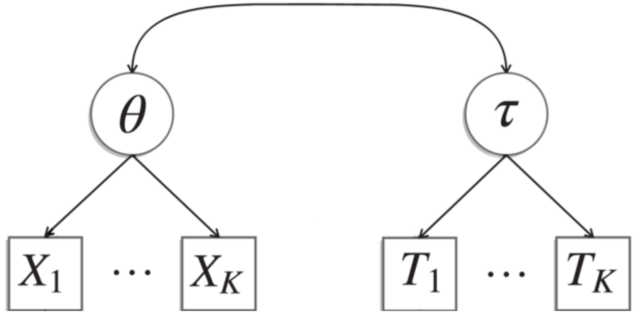
$$E(DT) = \frac{a}{2v}\frac{1 - e^{-av}}{1 + e^{-av}}$$

for reasonably high values of av (P+ close to 0/1)

$$\approx \frac{a}{2v} = \frac{1}{2}\frac{a_k^p/a_j^i}{v_k^p/v_j^i} = \frac{1}{2}\frac{v_j^i/a_j^i}{v_k^p/a_k^p}$$

$$E(\ln(DT)) = \xi_j - \tau_k \approx -\ln(2) + \ln\frac{v_j^i}{a_j^i} - \ln\frac{v_k^p}{a_k^p}$$

$$\xi_j \approx \ln\frac{v_j^i}{a_j^i};\ \tau_k \approx \ln\frac{v_k^p}{a_k^p}$$



$$\theta_k = a_k^p v_k^p \qquad \tau_k \approx \ln\frac{v_k^p}{a_k^p}$$

- ✓ positive correlation: primarily due to differences in v_k^p
- ✓ negative correlation: primarily due to differences in a_k^p

Dimensional analysis leads to sound results

$$E(DT) = \frac{a}{2v}\frac{1 - e^{-av}}{1 + e^{-av}}$$

for reasonably high values of av (P+ close to 0/1)

$$\approx \frac{a}{2v}$$

units of information

units information per seconds → measured in seconds

$$= \frac{1}{2}\frac{a_k^p/a_j^i}{v_k^p/v_j^i}$$

response caution: dimensionless quantity / time pressure: units of information

speed measure

What we have done

- a causal mechanism for the item response model: the Q-diffusion model
  - ✓ ability scale with natural zero point
  - ✓ incorporate guessing as part of the decision process
  - ✓ incorporate difficulty into discrimination parameter
  - ✓ find relation to the hierarchical model

Interesting avenues for further research

- a conjunctive multidimensional or multicomponent Q-diffusion model
- integrate these the cognitive diagnostic model and the Q-diffusion model through the construction of hierarchical models
- formulate the response time in the multiple-choice situation

Thank you for listening!