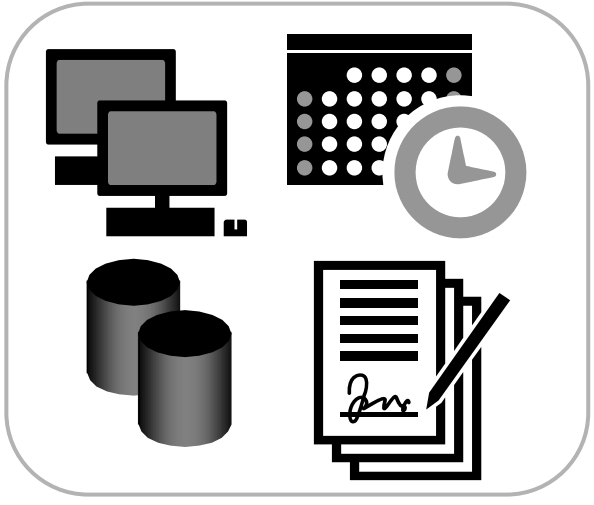
# Latent Trait Analysis for Risk Management of Complex Information Technology Projects

Speaker: T. Idé (Ide-san)
IBM T. J. Watson Research Center

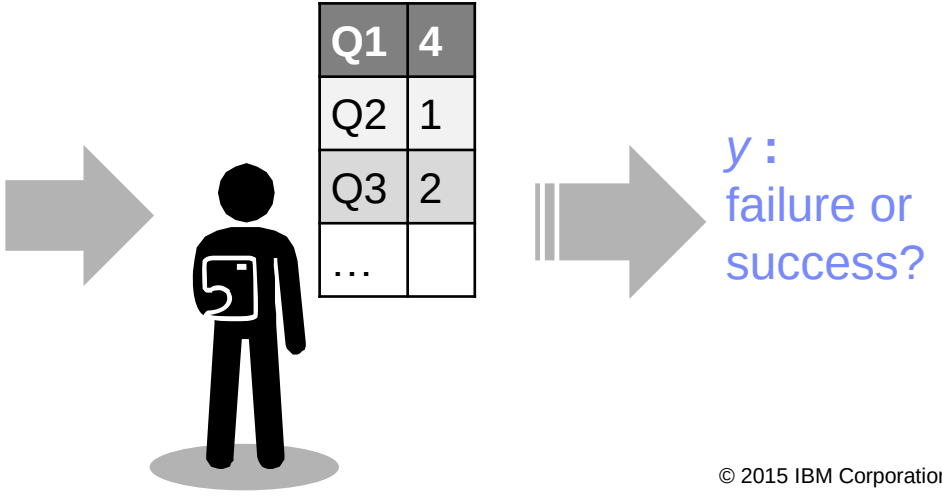
# Problem: Predict how much likely a project is going to fail

- **Input data, x : questionnaire answers**
  - Surveyor asks about the project status
  - Project manager answers to the questions
- **Predicted value,** y **: failure or success** (after contract signing)

IT system development project



***x*** : questionnaire answers

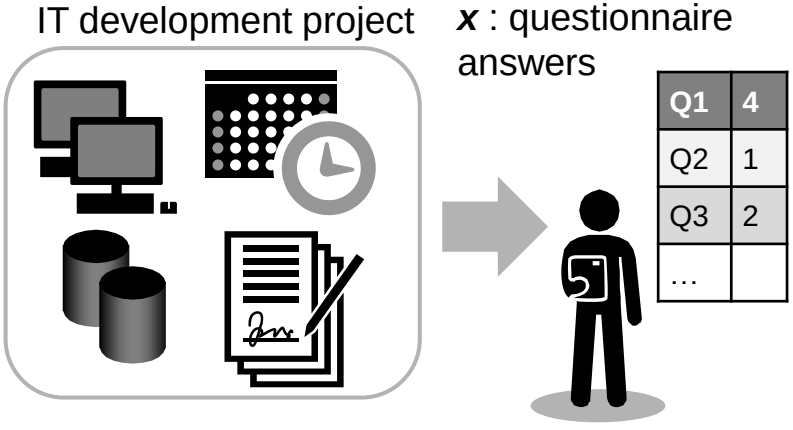



*y* : failure or success?

# (For ref.) What the questionnaire looks like

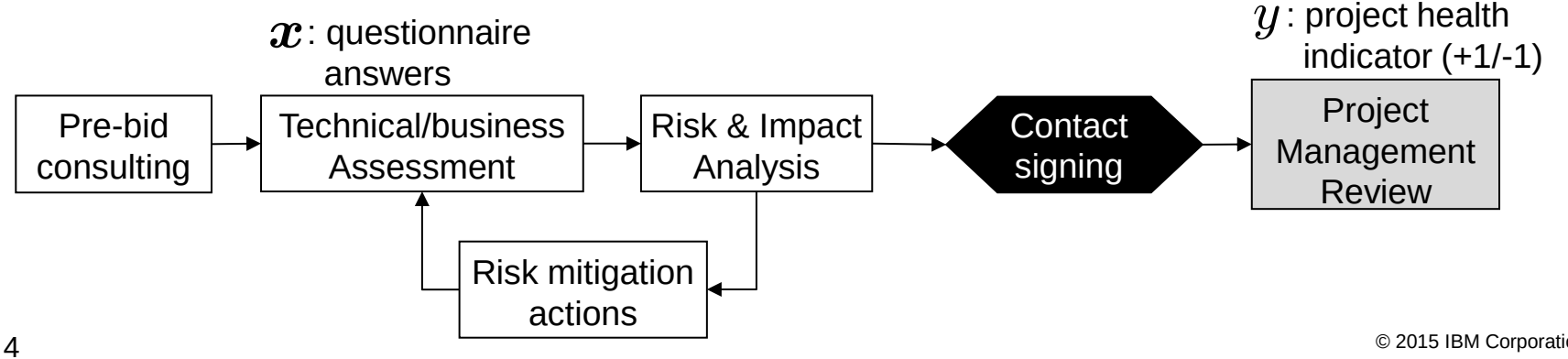
## □Major topics covered

- –Communication issues with the client
- –Well-definedness of the project scope
- –Issues related to subcontractors and internal teams
- –Project management issues



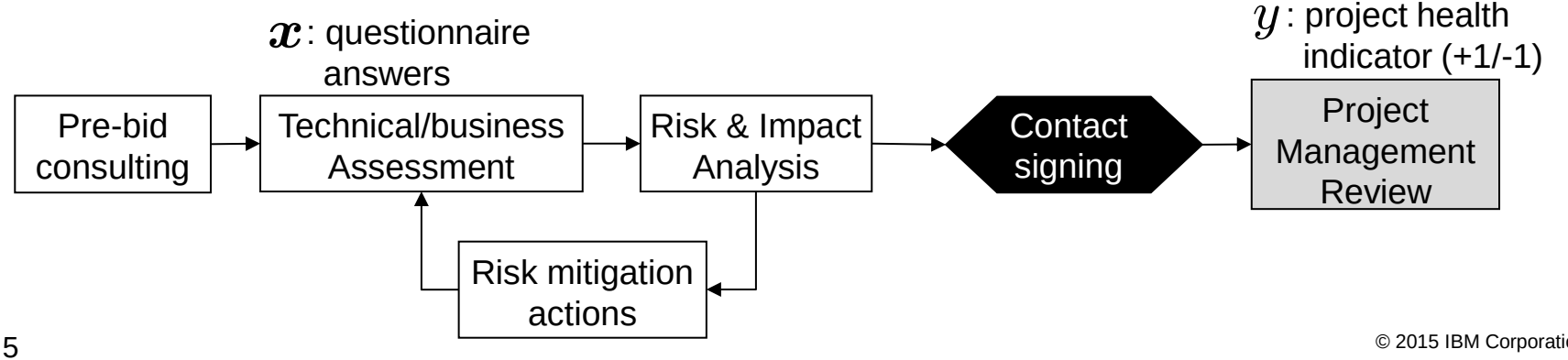
# Challenge (1): No evidently bad answers. Need to discover indications of failures from apparently good answers

- **Iterative review process allows removing all evident risk factors**
  - This is actually a prerequisite to get into the final review right before contract signing
- **However, some of them might be "pretending" as good**
- **Wish to discover such indications**



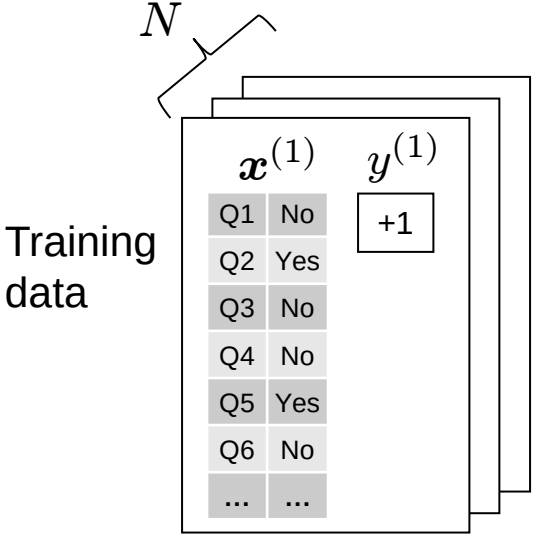
# Challenge (2): No systematic way of evaluating the goodness of questionnaire items

- **Individual questions are defined by human experts based on their experience**
- **Some questions may be unnecessary and/or incur too much cost to answer**
- **Wish to have a quantitative approach to evaluate the *informativeness* of each questions.**



# Problem summary

- **1. Compute the informativeness of the question items**
- **2. Build a predictive model for project failure/success (*y*) given a new set of questionnaire answers (*x*)**




$$\boldsymbol{x} \in \{0,1\}^M$$

M questions
- 1: at-risk
- 0: no risk

Prediction model

| | |
|---|---|
| Q1 | Yes |
| Q2 | No |
| Q3 | No |
| Q4 | No |
| Q5 | No |
| Q6 | Yes |
| ... | ... |

?

$$y \in \{-1,+1\}$$

# Key idea: Assume ***x*** is stochastically generated by a latent variable that is more faithful to the truth

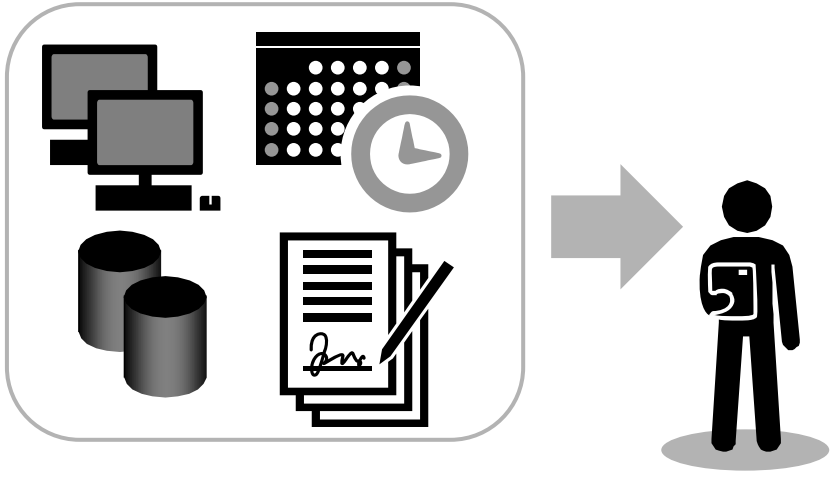
**latent failure tendency**

$$\theta \rightarrow x$$

questionnaire answer

**Should be more faithful to the truth**

**May be disguised as good**



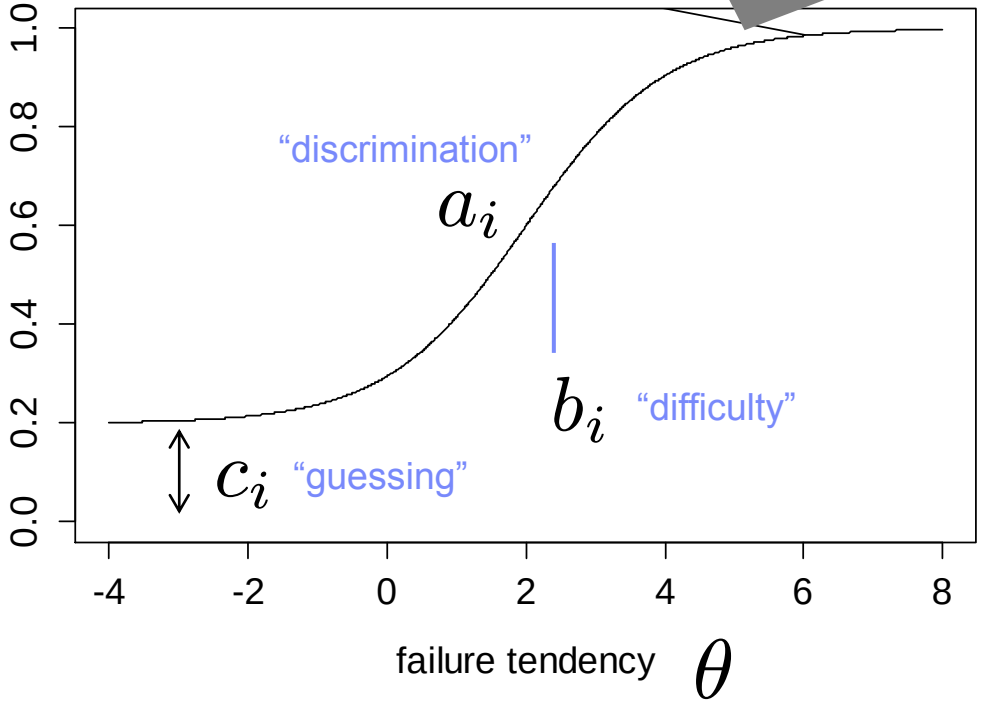
# Using a “shifted S-curve” as a natural model of cognitive biases

- Overly optimistic for smaller risks
- Overly pessimistic for larger risks
- Sometimes use a guess

□**Represents nonlinear relationship of** $\theta \rightarrow x$

Prob. of answering as at-risk for the *i*-th question

$P(\theta, a_i, b_i, c_i)$




This type of model is known as **latent trait model** in psychometrics

# (For ref.) Latent trait model

- **The method to normalize the SAT test**
  - The total score is NOT simply the total number of accurately answered items
    - Scores are normalized to be in the range of [200-800].
    - Even random guess to all of the items gives you at least 200 points to each section.
  - Examinee's ability is treated as a latent variable to be estimated
- **In our case, examinees' ability corresponds to latent failure tendency**

# Bayesian framework to estimate latent failure tendency

Historical record

$$\left(\boldsymbol{x}^{(n)}, y^{(n)}\right)$$

$$n = 1, \ldots, N$$

**training phase**

Generative model for x
based on the shifted S-curve

$$p(\boldsymbol{x} \mid \theta, \boldsymbol{a}, \boldsymbol{b}, \boldsymbol{c})$$

model
parameters

$$\hat{\boldsymbol{a}}, \hat{\boldsymbol{b}}, \hat{\boldsymbol{c}}$$

New question
answer $\boldsymbol{x}$

**prediction phase**

Prior distribution for theta

$$f(\theta) = \sqrt{\frac{1}{2}} \exp\left(-\frac{1}{2}\theta^2\right)$$

estimated
failure
tendency $\hat{\theta}$

# Making prediction using estimated latent failure tendency

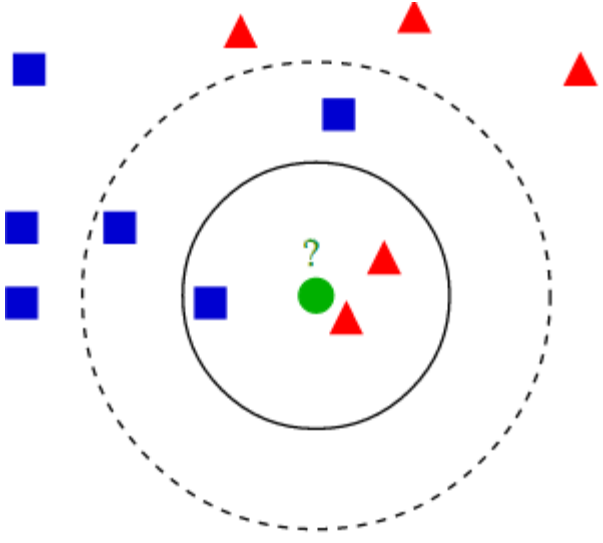
New question answer $\boldsymbol{x}$

**prediction phase**

estimated failure tendency $\hat{\theta}$

Simply use k-NN classification
based on the distance in the latent space

$$d(\boldsymbol{x}, \boldsymbol{x}^{(n)}) = \sum_{g=1}^{G} w_g \left( \hat{\theta}_g - \hat{\theta}_g^{(n)} \right)^2$$



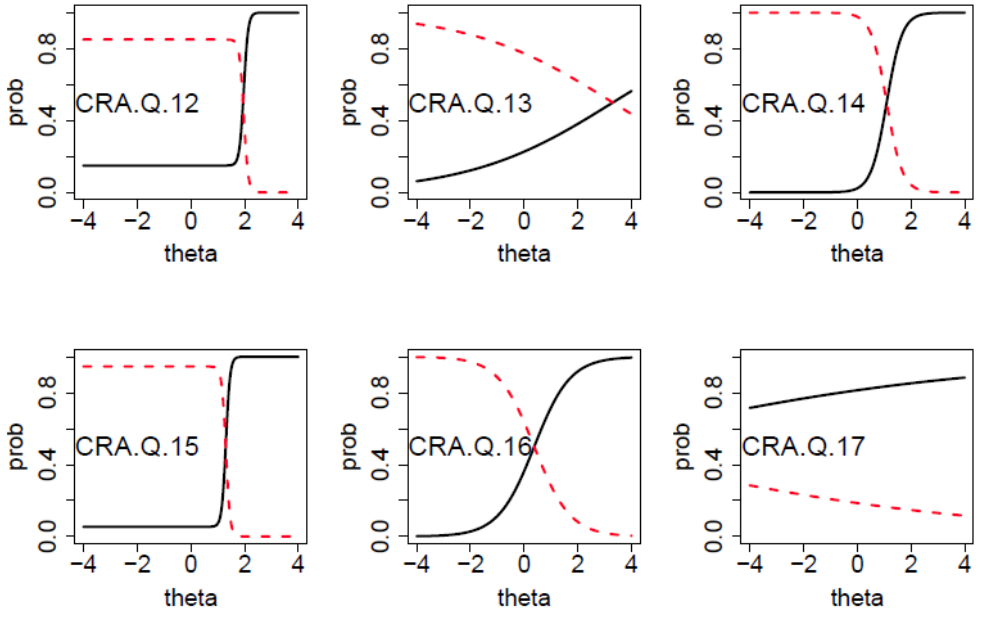

Extend the original LTA to include multiple latent variables (→ see the paper)

# Experiment: Using service provider's real quality assurance data

- **Questionnaire called CRA (contract risk assessment)**
- **M = 22 rather qualitative questions**
- **N = several hundred**
- **Each question is yes (at-risk) or no (no-risk)**
- **Final project evaluation is failure (y=+1) or non-failure (y= -1)**

# Result (1): Estimated S-curves clearly provide practical information on the usefulness of each question

- **Q17 is hardly useful**
  - Formal question on service pricing
  - Expected to be less influential
- **Q14 and Q16 are useful**
  - Ask straight about the clarity and feasibility of the development plan



# Result (2): Achieved better accuracy

- **Compares F-value**
  - harmonic mean between troubled project accuracy and non-troubled project accuracy
- **Clearly outperform the baseline**
  - Baseline is based only on **x**
    - Logistic regression
    - Simple k-NN
  - Our approach uses theta instead of **x**

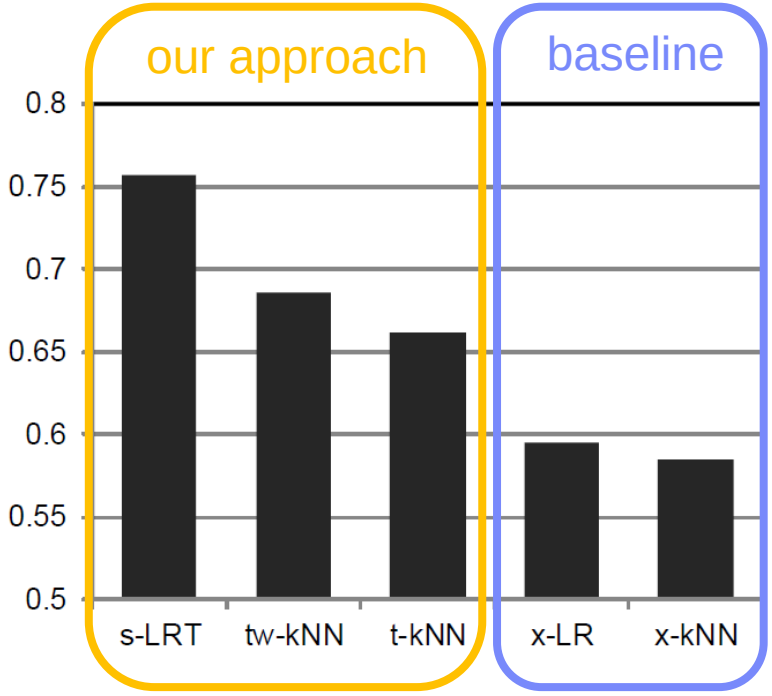



Fig. 8. Comparison of failure prediction accuracies.

# Conclusion:

- **Proposed a new approach to project risk management inspired by a psychometric theory**

- **By correcting cognitive biases, we confirmed that the proposed method**
  - provides practically useful information on the usefulness of each question
  - clearly outperforms known alternatives

Thank you!

# Our session TS9: Big Data for Management, is scheduled for Wed, May 13, 2015, 14:00 - 15:30

- **The order of presentation is:**
  - Early Network Failure Detection System by Analyzing Twitter Data
    - Kei Takeshita, Masahiro Yokota, Ken Nishimatsu, NTT, Japan.
  - Resolution Recommendation for Event Tickets in Service Management
    - Wubai Zhou, Liang Tang, Tao Li, Florida International University, USA,
    - Larisa Shwartz, IBM T.J. Watson Research Center, USA,
    - Genady Grabarnik, St. John's University, USA.
  - Dude, Ask The Experts!: Android Resource Access Permission Recommendation with RecDroid
    - Bahman Rashidi, Carol Fung, Virginia Commonwealth University, Canada,
    - Tam Vu, University of Colorado Denver, USA.
  - Latent Trait Analysis for Risk Management of Complex Information Technology Projects
    - Tsuyoshi Ide, Sinem Guven, Ea-Ee Jan, IBM T. J. Watson Research Center, USA,
    - Sergey Makogon, Alejandro Venegas, IBM Global Technology Services, Algeria.

- **Chair: Prof Hanan Lutfiyya**
  - ハナン・ルトフィヤ教授
  - presentation should last around 17 minutes within the 22 minutes allocated slot time to leave at least (!) 5 minutes for questions.