



Verso una misura oggettiva: criteri per la realizzazione e la valutazione di test a risposta aperta

Daniele Vidoni

CRELL – Commissione Europea

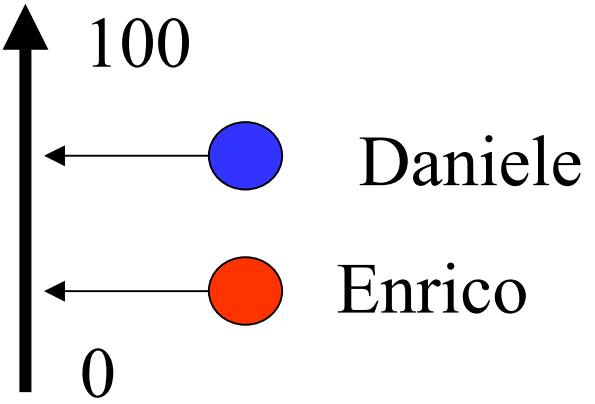
daniele.vidoni@jrc.it

La valutazione

Joint Research Centre

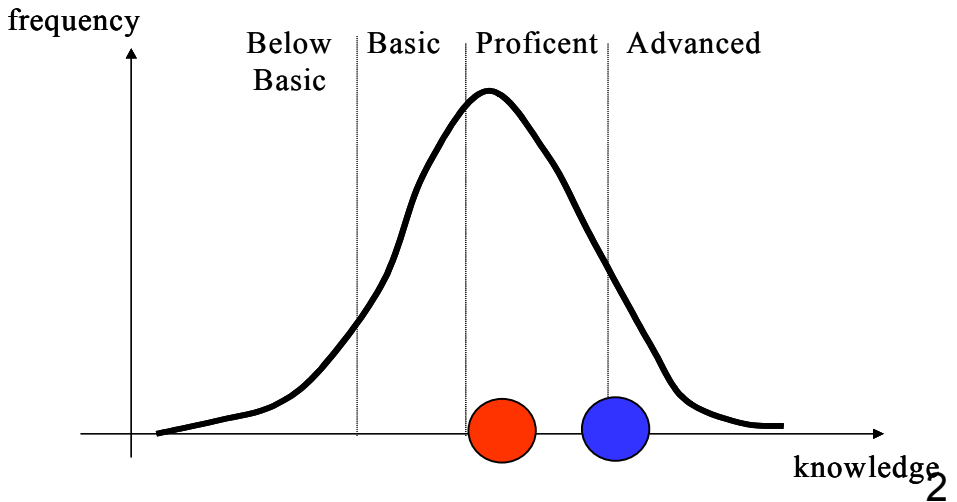
Criterio

% di risposte esatte



Norma

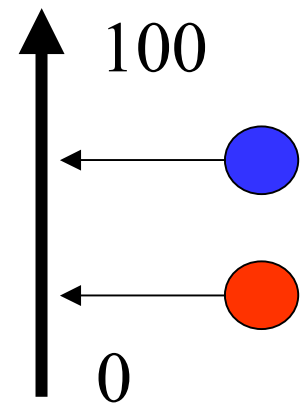
Livello raggiunto rispetto agli altri studenti



Joint Research Centre

Criterion

% di risposte esatte
nel test

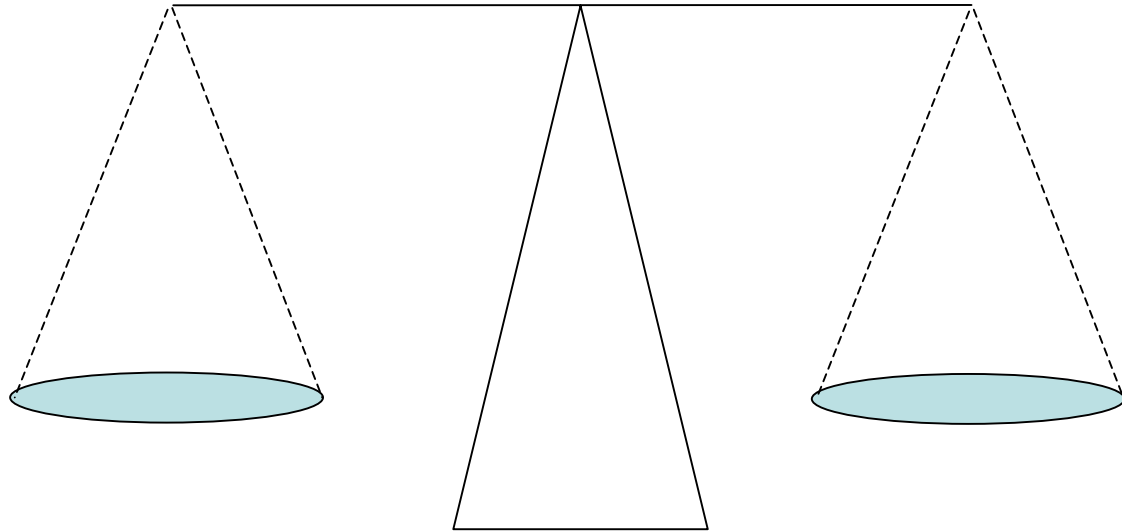
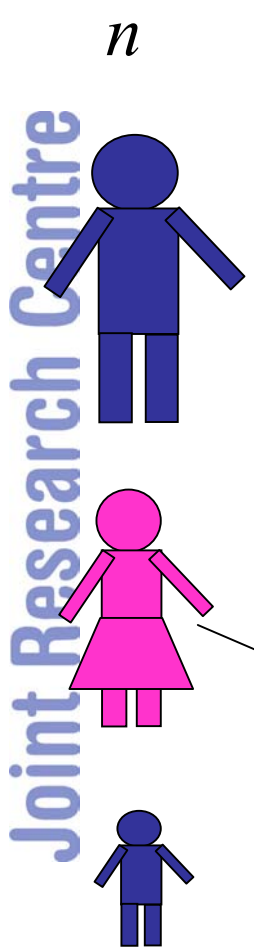


Problemi:

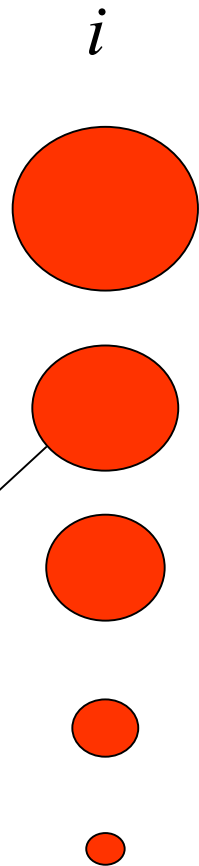
I risultati dipendono dalla
difficoltà delle domande

Esempio:

Individuos



Piedras



Cerchiamo il peso dei bambini e delle pietre

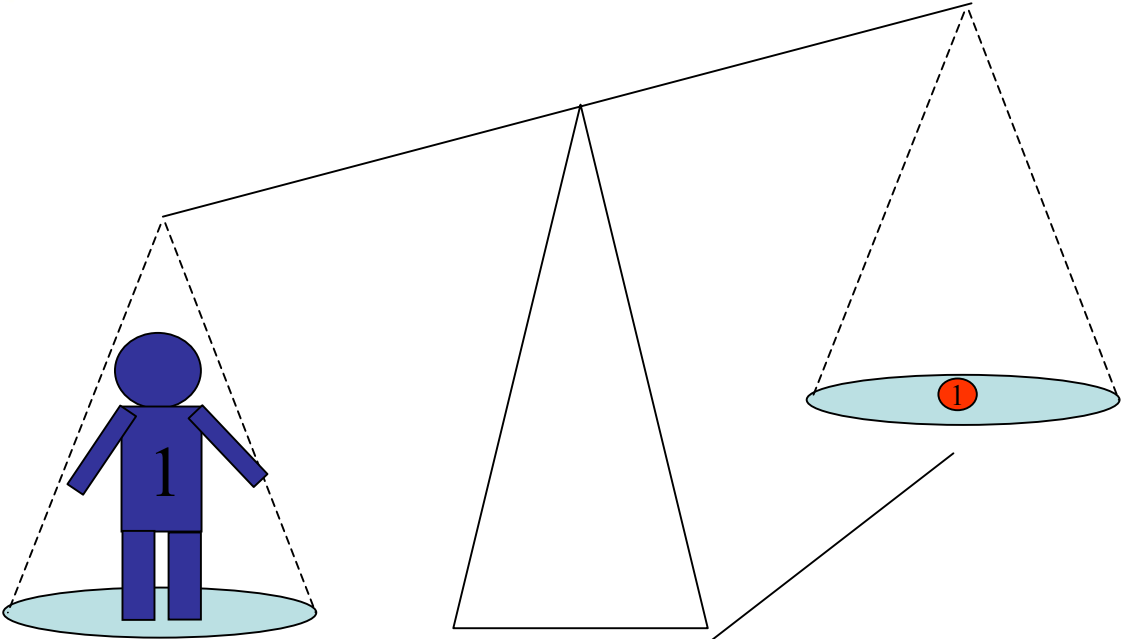
- pesante

+ pesante

Individui

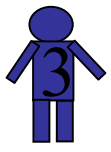
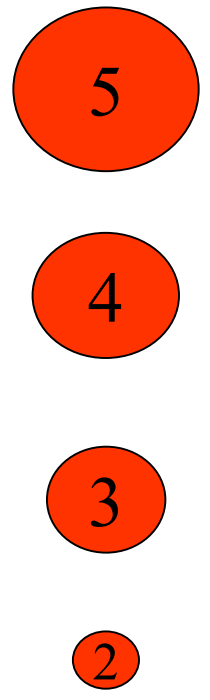
Joint Research Centre

n



Pietre

i



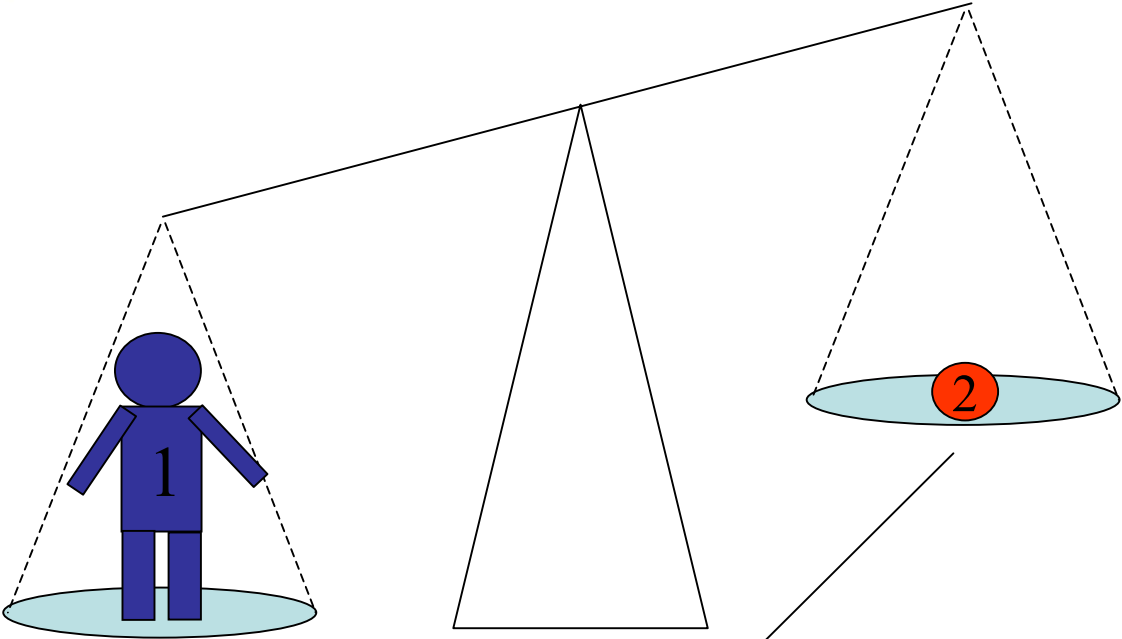
$n \setminus i$	1	2	3	4	5
1	1				
2					
3					

1 = "supera la prueba"

Individui

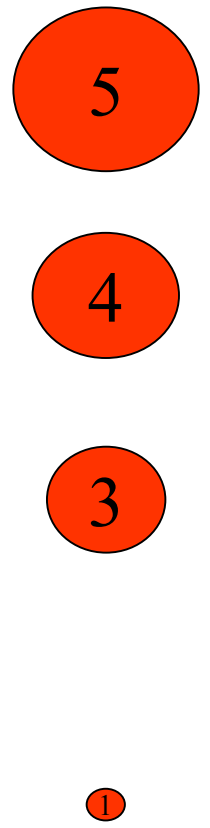
Joint Research Centre

n



Pietre

i



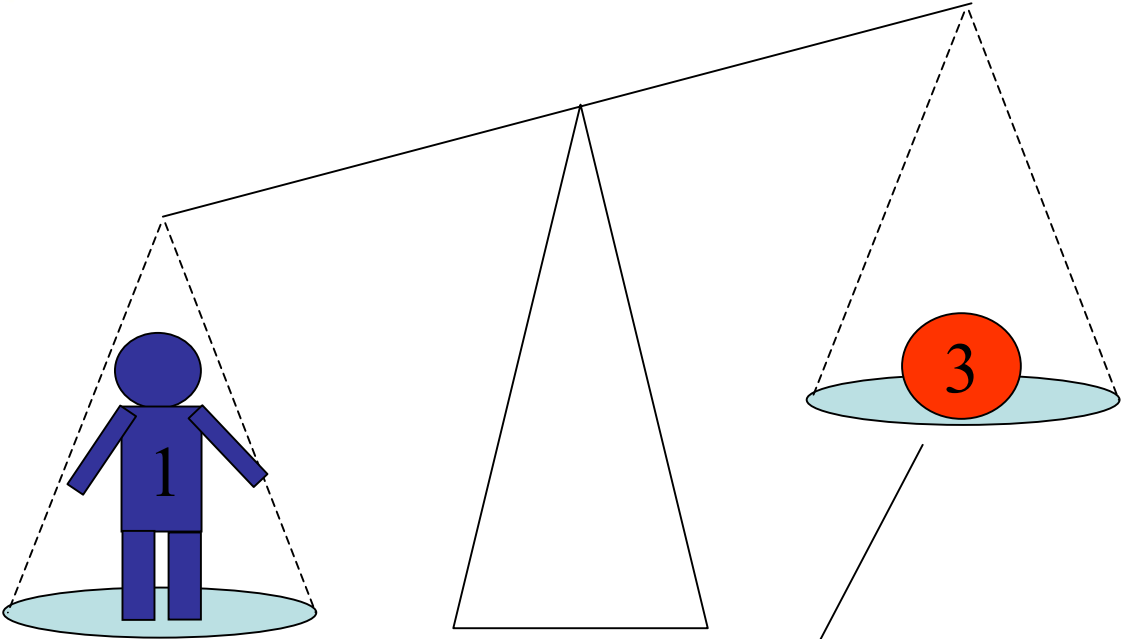
$n \setminus i$	1	2	3	4	5
1	1	1			
2					
3					

1 = "supera la prueba"

Individui

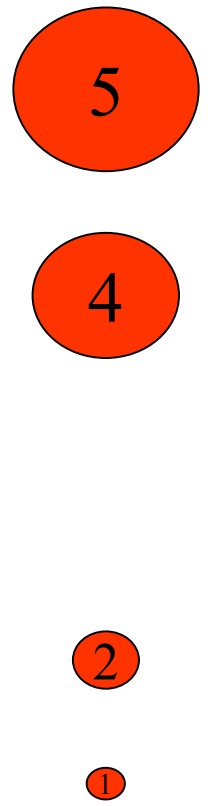
Joint Research Centre

n



Pietre

i



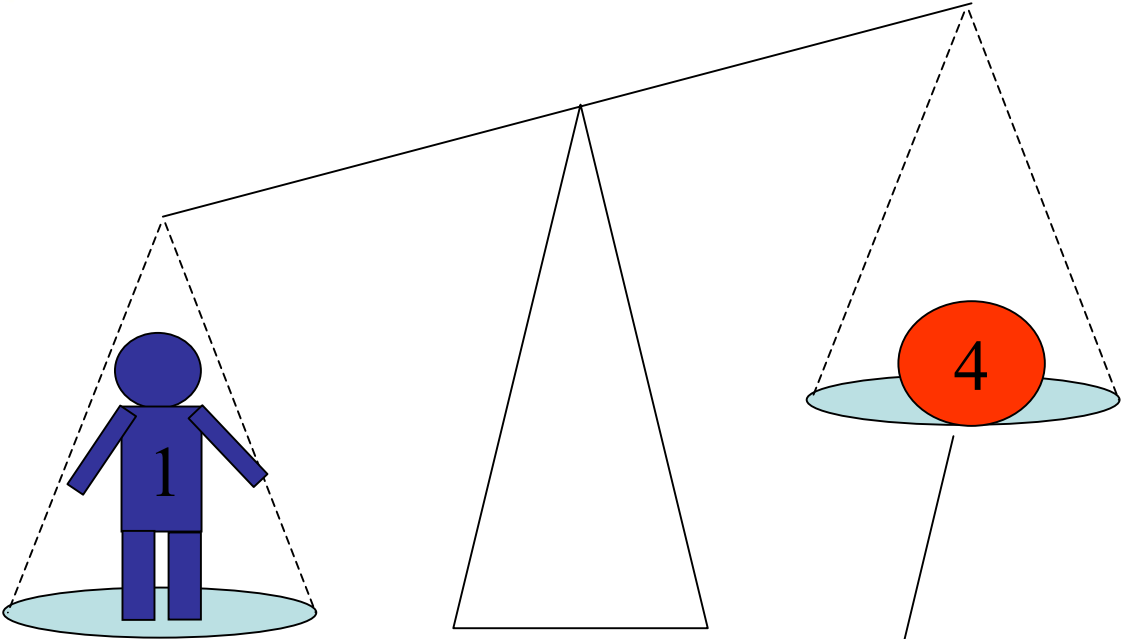
$n \setminus i$	1	2	3	4	5
1	1	1	1		
2					
3					

1 = “supera la prueba”

Individui

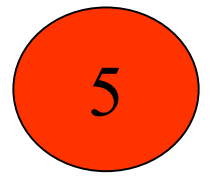
Joint Research Centre

n



Pietre

i



$n \setminus i$	1	2	3	4	5
1	1	1	1	1	
2					
3					

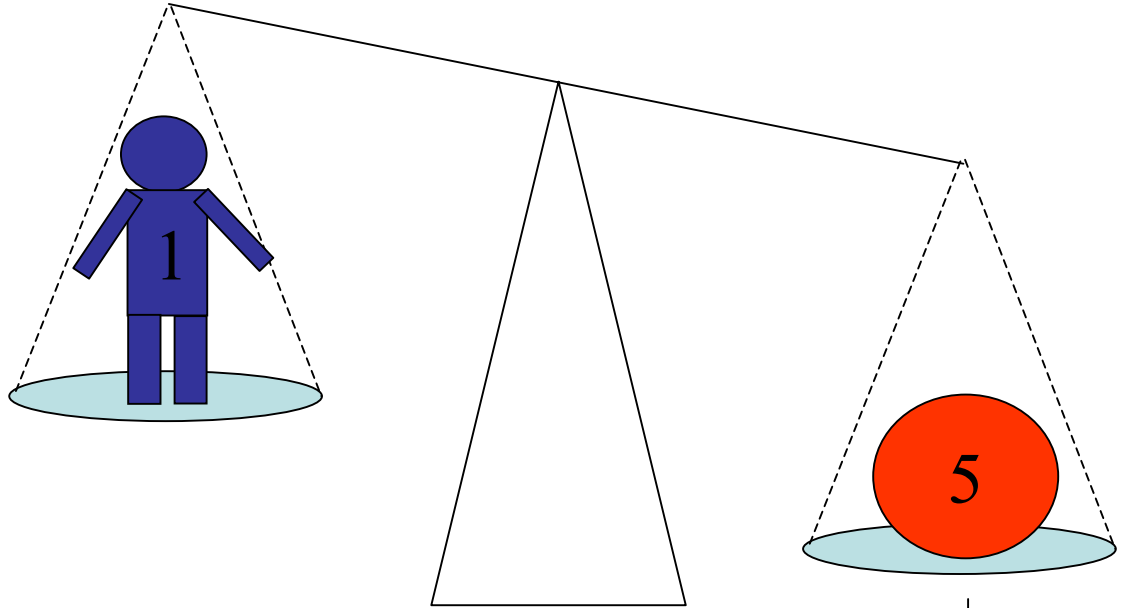
1 = "supera la prueba"

Individui

Pietre

n

i



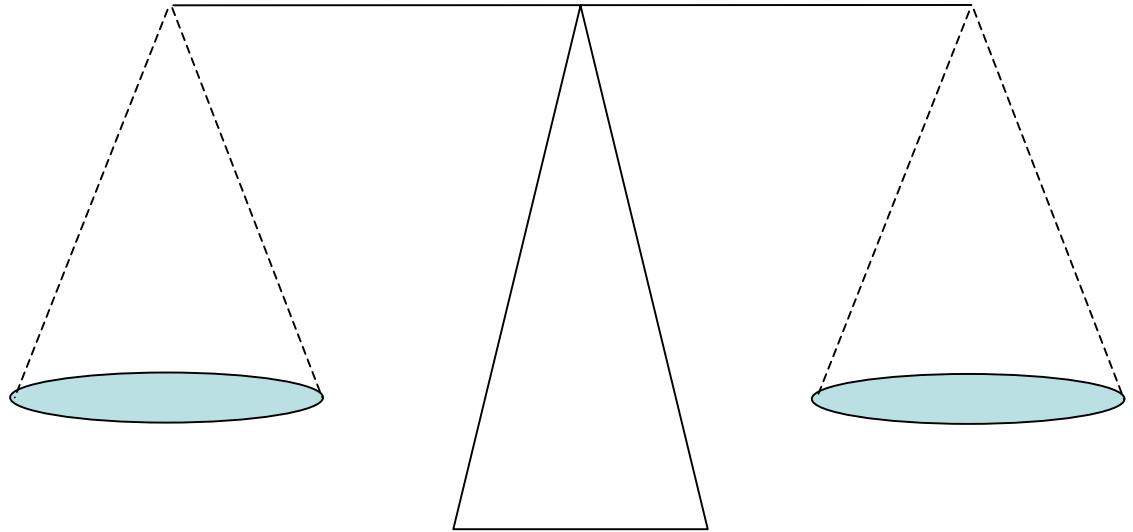
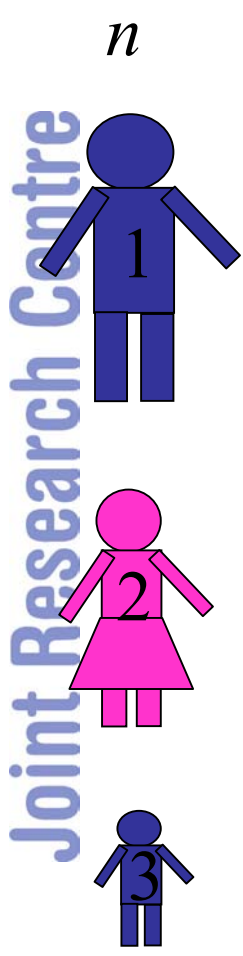
Joint Research Centre



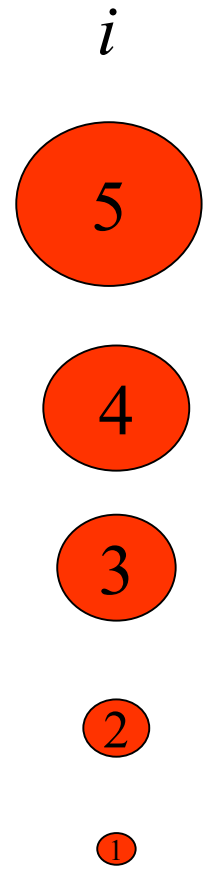
$n \setminus i$	1	2	3	4	5
1	1	1	1	1	0
2					
3					

0 = no "supera la prueba"

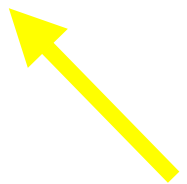
Individui



Pietre



$n i$	1	2	3	4	5	Score
1	1	1	1	1	0	4
2	1	1	1	0	0	3
3	1	1	0	0	0	2



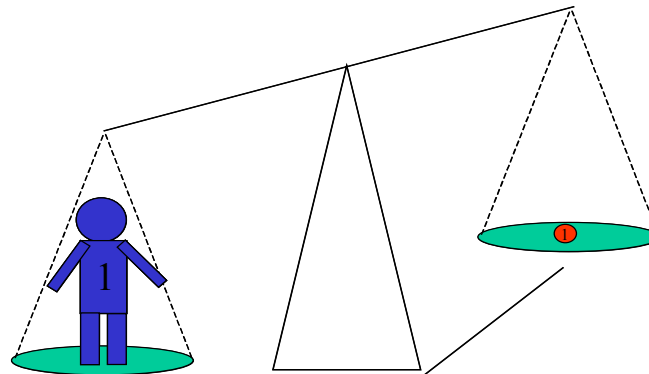
Medida de criterio ¹⁰

Joint Research Centre

Non **pesa** di più chi supera il maggior numero di prove, bensì chi supera la proba più **“pesante”**

individui

..



pietre



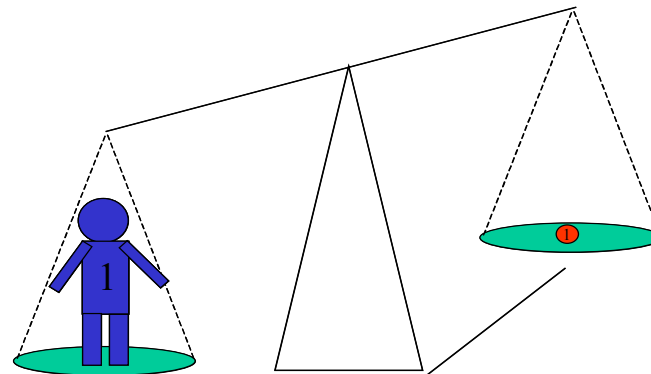
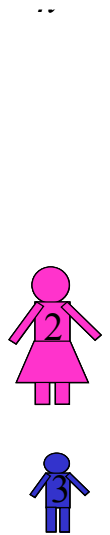
$n \setminus i$	1	2	3	4	5
1	1				
2					
3					

1 = supera la proba

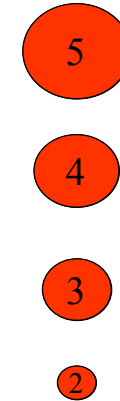


Non è più **abile** chi risponde al maggior numero di domande, bensì chi risponde alle domande più “**difficili**”

individui



items



$n \backslash i$	1	2	3	4	5
1	1				
2					
3					

1 = supera la prova



Per fare confronti è
necessario misurare le
differenze tra i livelli di
conoscenza degli studenti



Misure standardizzate

- Le differenze si possono misurare solo utilizzando misure standard attraverso **unità iterative**.
- Ciò permette di misurare la **crescita** (differenze nel tempo).



Georg Rasch (1901-1980)

Probabilistic Models for Some Intelligence and Attainment Tests

GEORG RASCH
WITH A FOREWORD AND
AFTERWORD BY
BENJAMIN D. WRIGHT

Mesa Press,
Chicago 1980

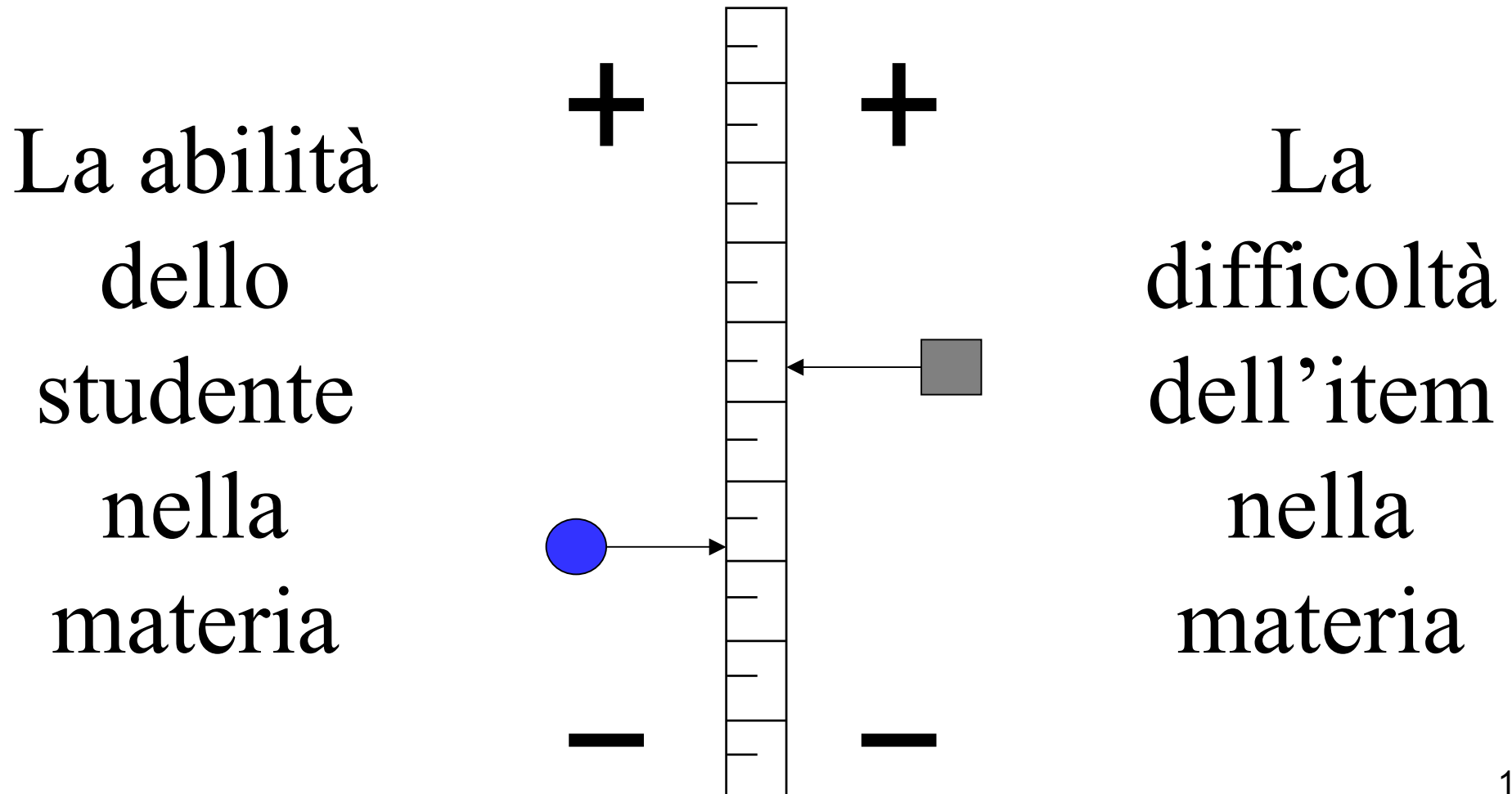


È necessario sviluppare misure
che siano indipendenti dalla
abilità della popolazione e dalla
difficoltà degli item

George Rasch (1960)

Specific objectivity

Nella teoria di Rasch la probabilità di rispondere correttamente ad un item deve dipendere solo da due fattori

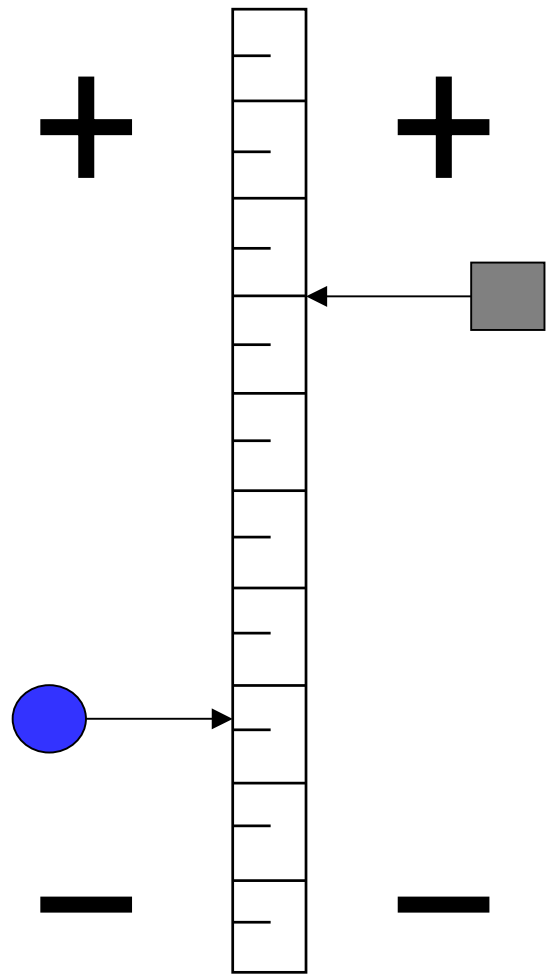


Probabilità di rispondere correttamente

Molto bassa ≈ 0

Abilità

Difficoltà

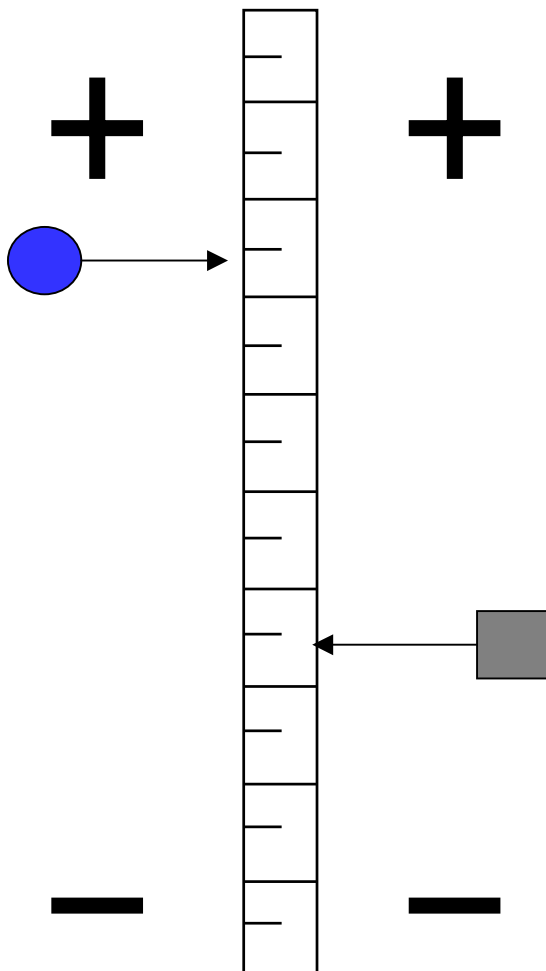


Probabilità di rispondere correttamente

Molto alta ≈ 1

Abilità

Difficoltà



Probabilità di rispondere correttamente

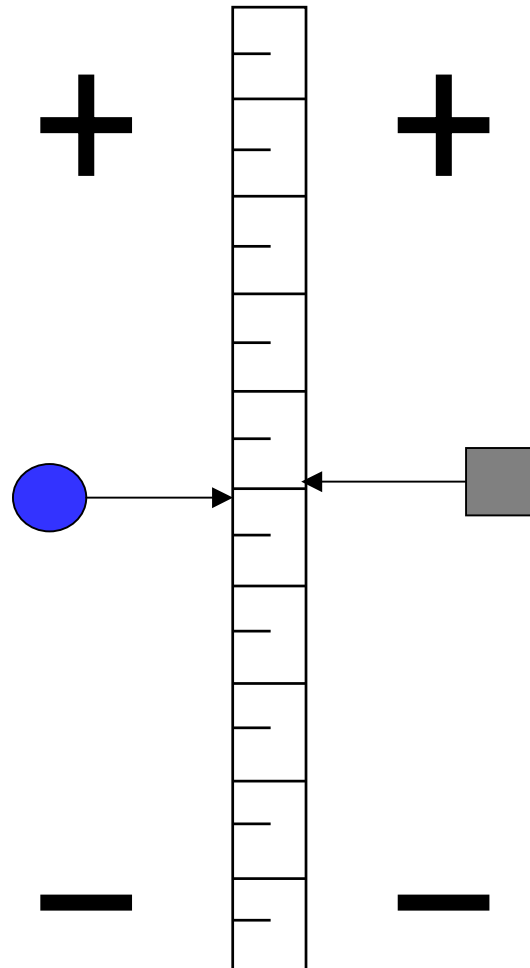
“Media” $\approx 0,5$

Abilità

+

+

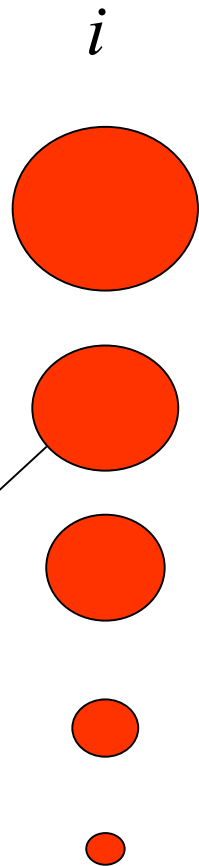
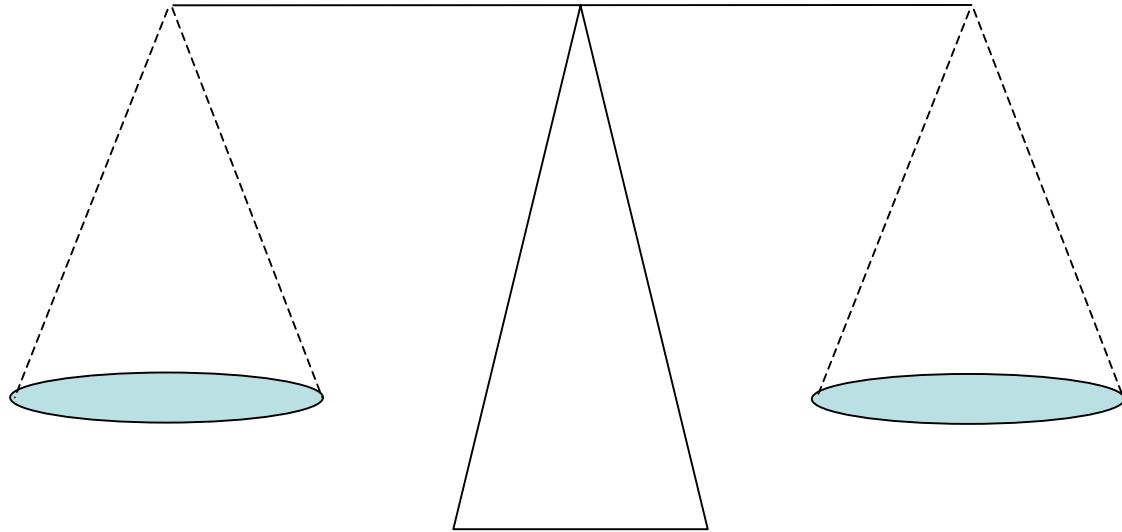
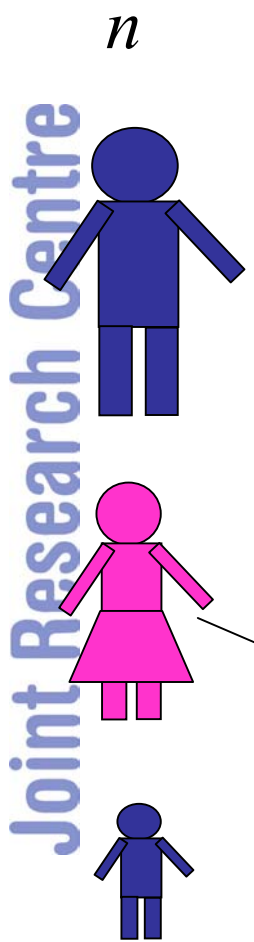
Difficoltà



La teoria di Guttman-Rasch

Individui

Item



Cerchiamo l'abilità degli individui e la difficoltà degli item

- pesante

+ pesante

La teoria di Guttman-Rasch

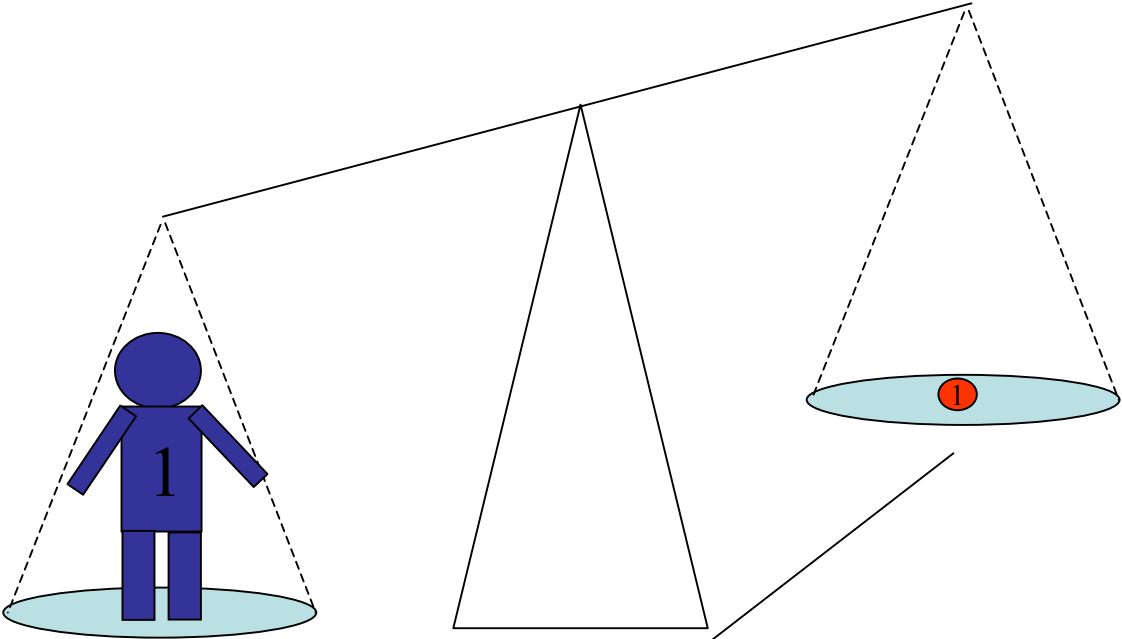
Individui

Items

n

i

Joint Research Centre



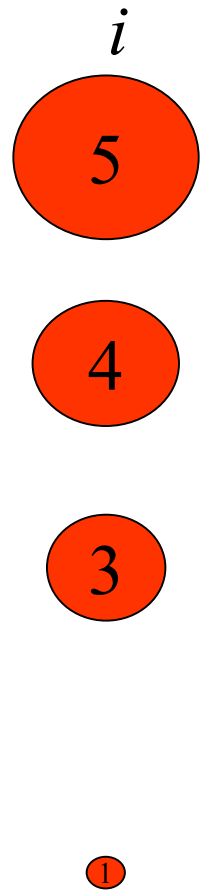
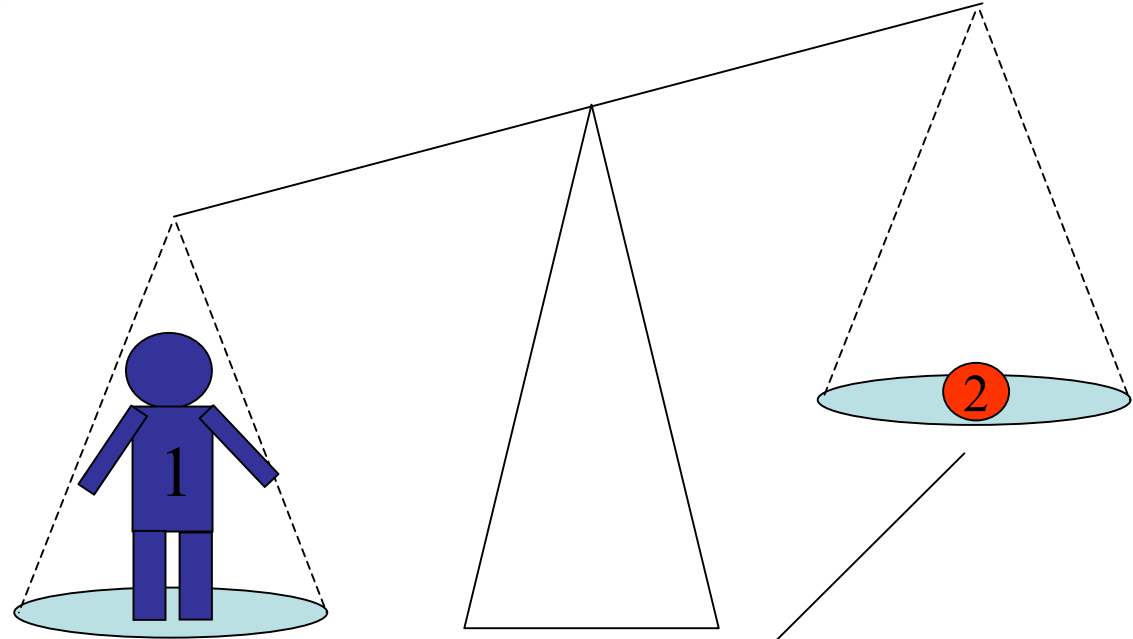
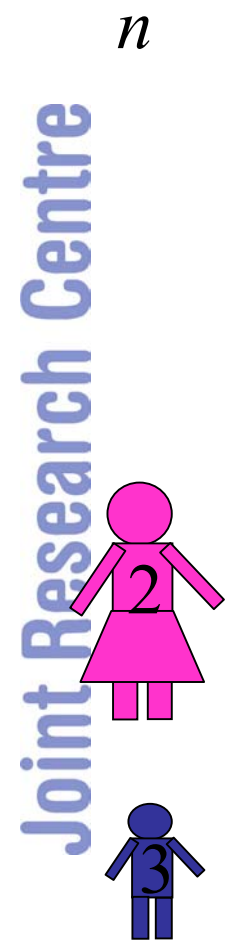
$n \setminus i$	1	2	3	4	5
1	1				
2					
3					

1 = “supera la prova”

La teoria di Guttman-Rasch

Individui

Items



$n \setminus i$	1	2	3	4	5
1	1	1			
2					
3					

1 = “supera la prova”

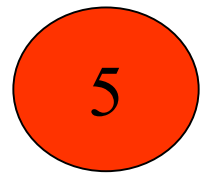
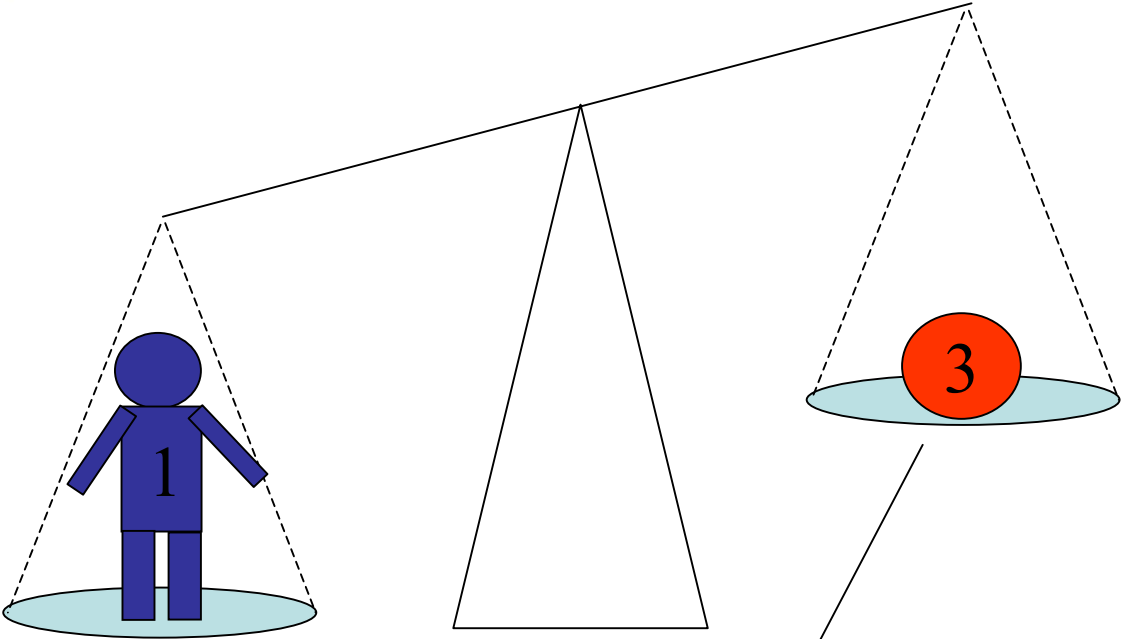
La teoria di Guttman-Rasch

Individui

Items

n

i



Joint Research Centre



$n \setminus i$	1	2	3	4	5
1	1	1	1		
2					
3					

1 = “supera la prova”

La teoria di Guttman-Rasch

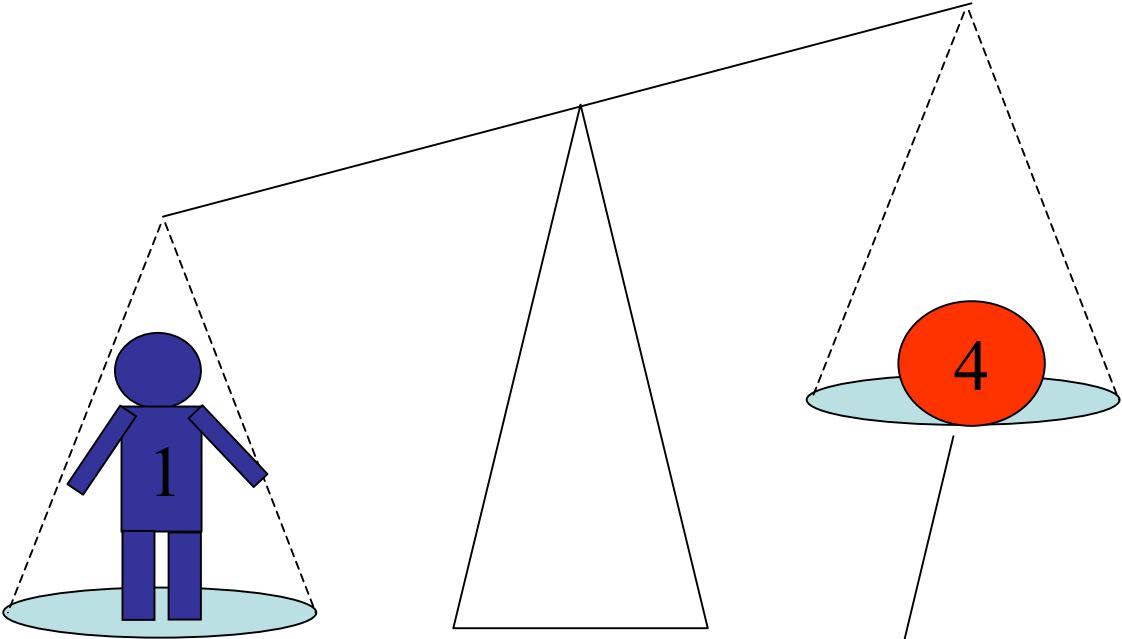
Individui

Items

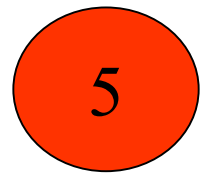
n

i

Joint Research Centre



$n \setminus i$	1	2	3	4	5
1	1	1	1	1	
2					
3					



1 = “supera la prova”

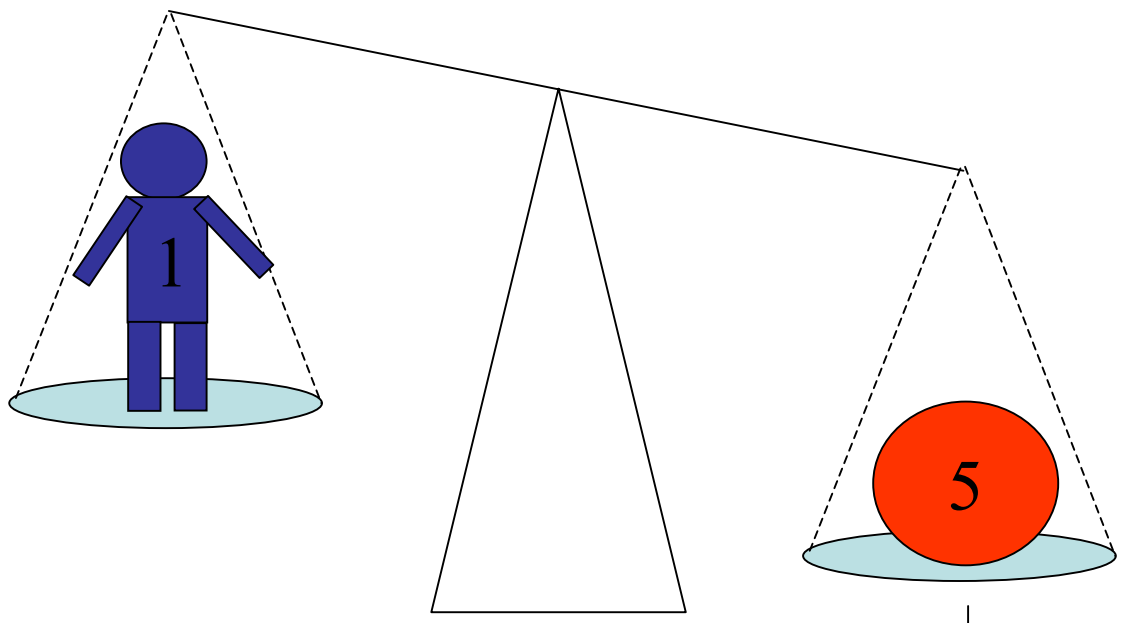
La teoria di Guttman-Rasch

Individui

items

n

i



Joint Research Centre



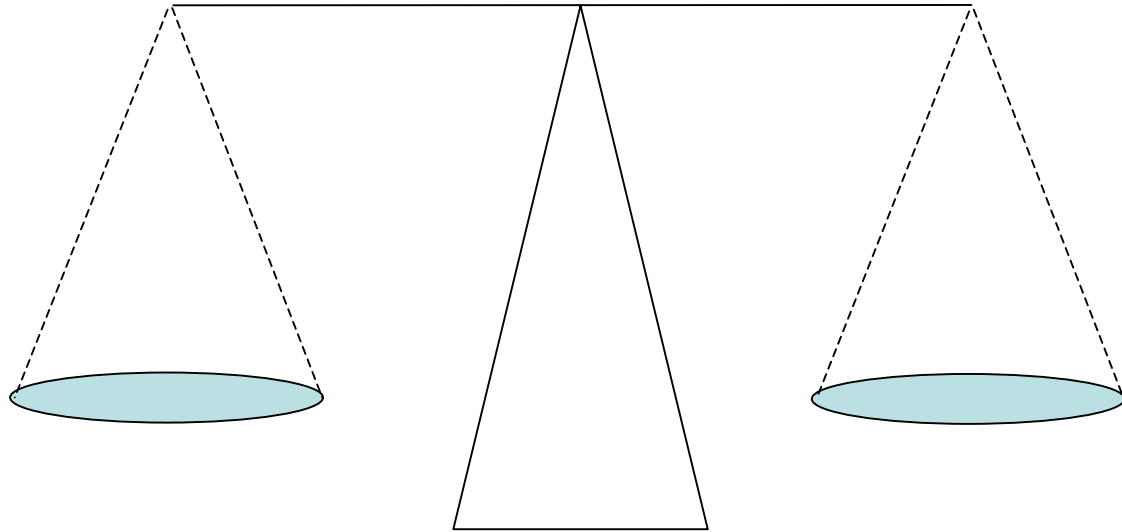
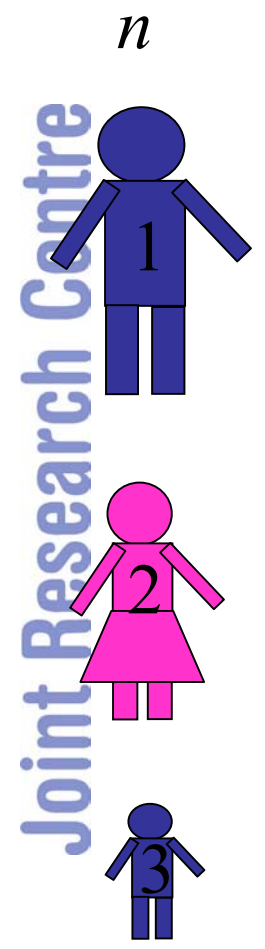
$n \setminus i$	1	2	3	4	5
1	1	1	1	1	0
2					
3					



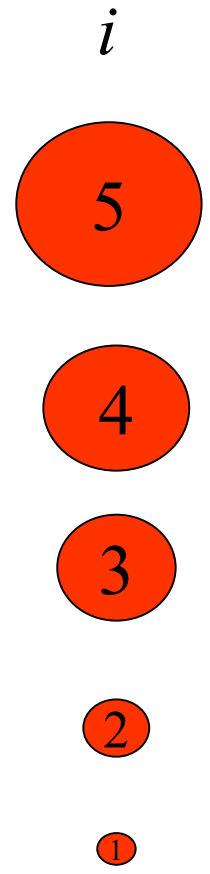
0 = non “supera la prova”

La teoria di Guttman-Rasch

Individui

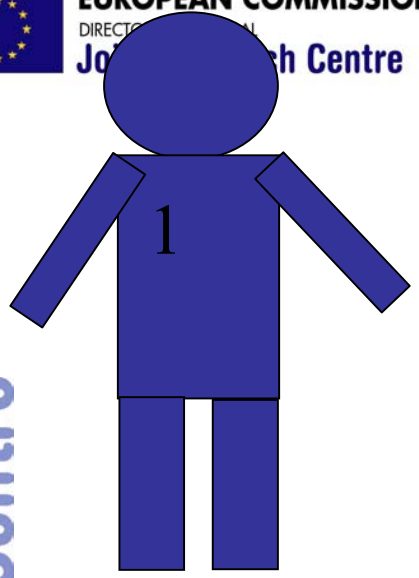


Items

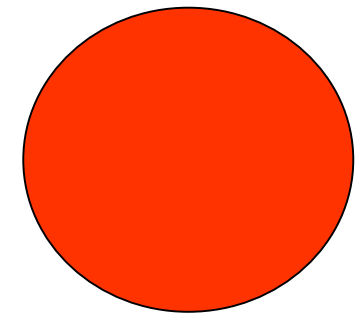


Dati per stimare la **abilità** e la **difficoltà**

$n \setminus i$	1	2	3	4	5
1	1	1	1	1	0
2	1	1	1	0	0
3	1	1	0	0	0

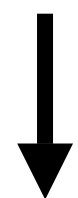


Anche se il peso è esattamente uguale

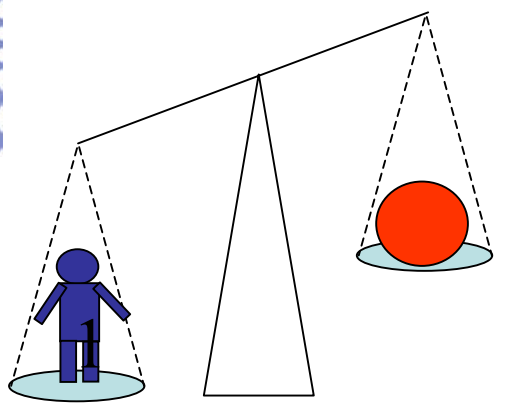


Joint Research Centre

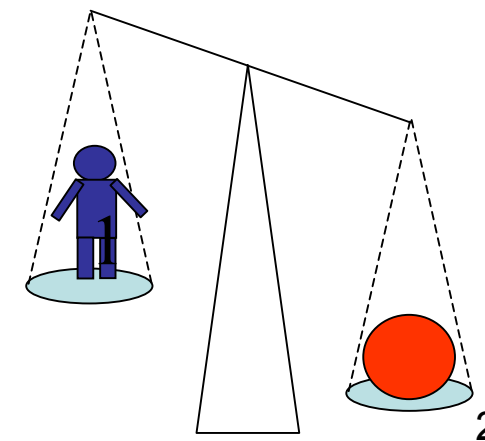
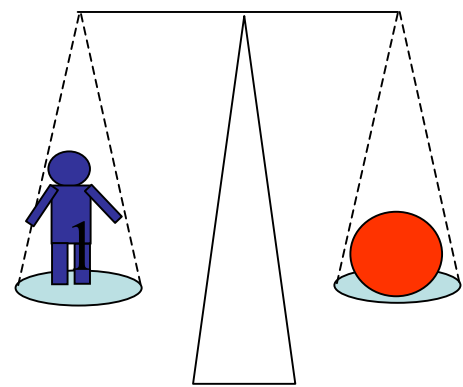
Non esiste una bilancia tanto precisa da poter osservare questa situazione



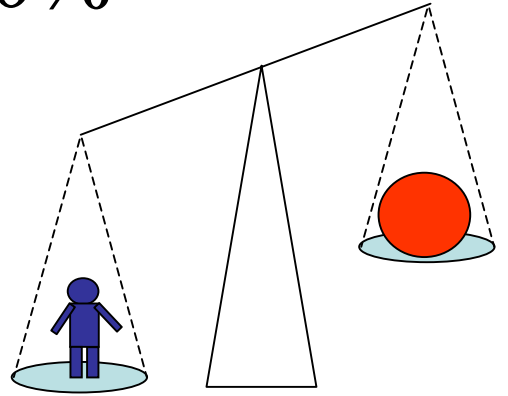
50%



50%

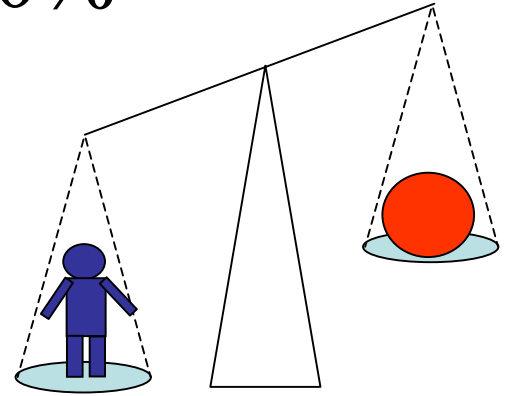


10%



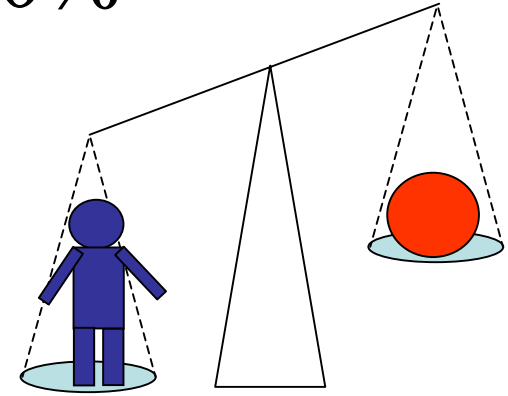
Risultato nella prova = 1

50%



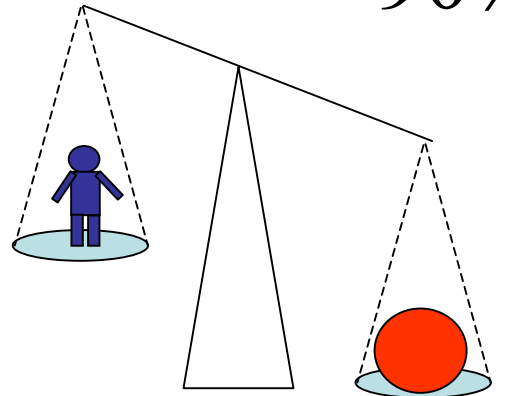
Risultato nella prova = 1

90%



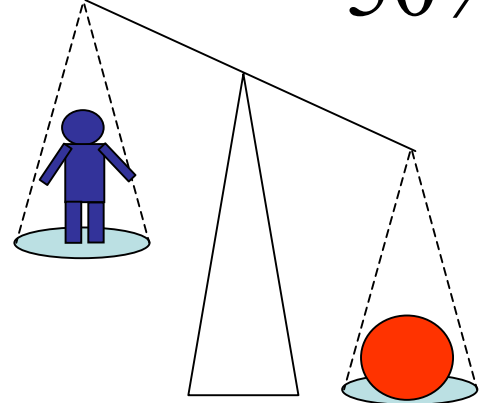
Risultato nella prova = 1

90%



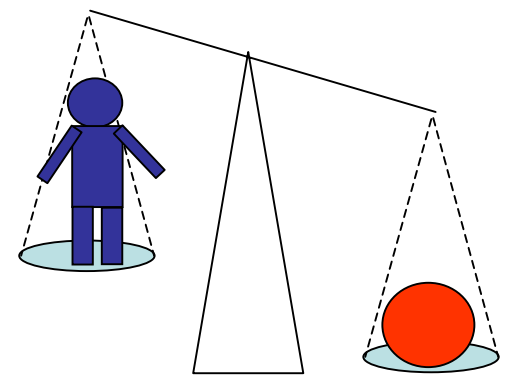
Risultato nella prova = 0

50%



Risultato nella prova = 0

10%

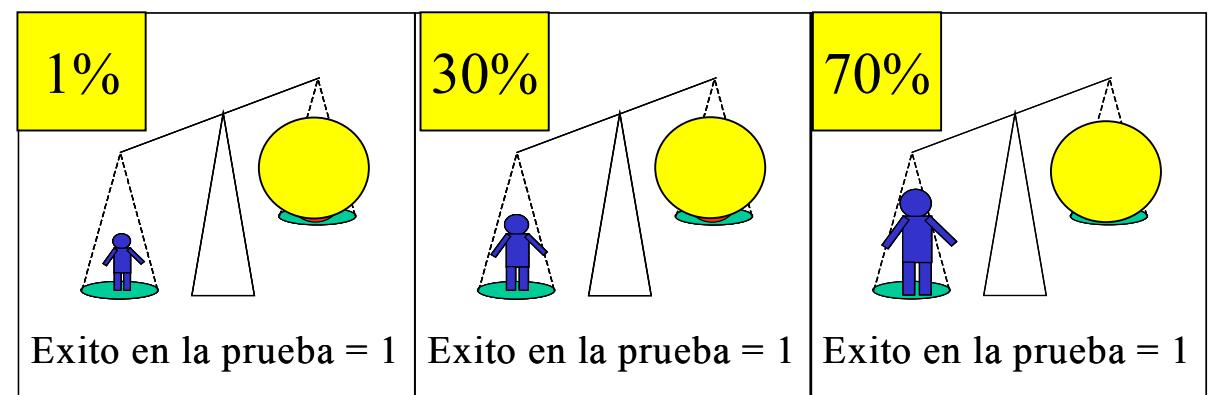


Risultato nella prova = $\frac{0}{29}$

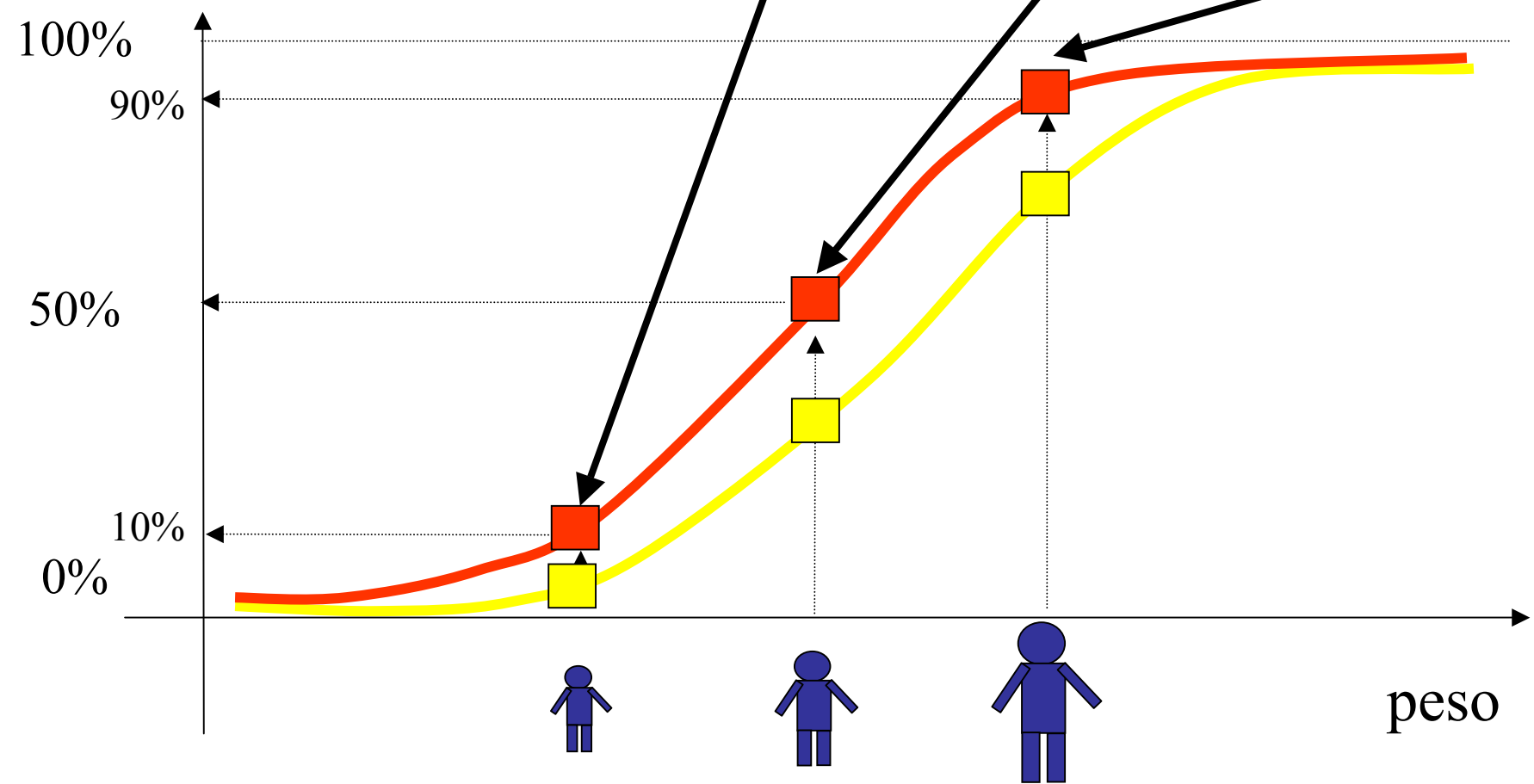


E se la pietra pesa di più?

Probabilità che il risultato sia = 1



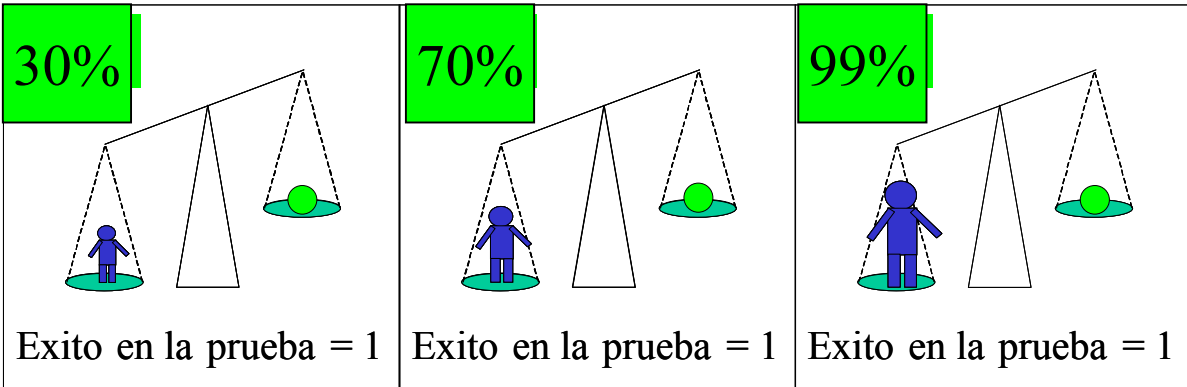
Joint Research Centre



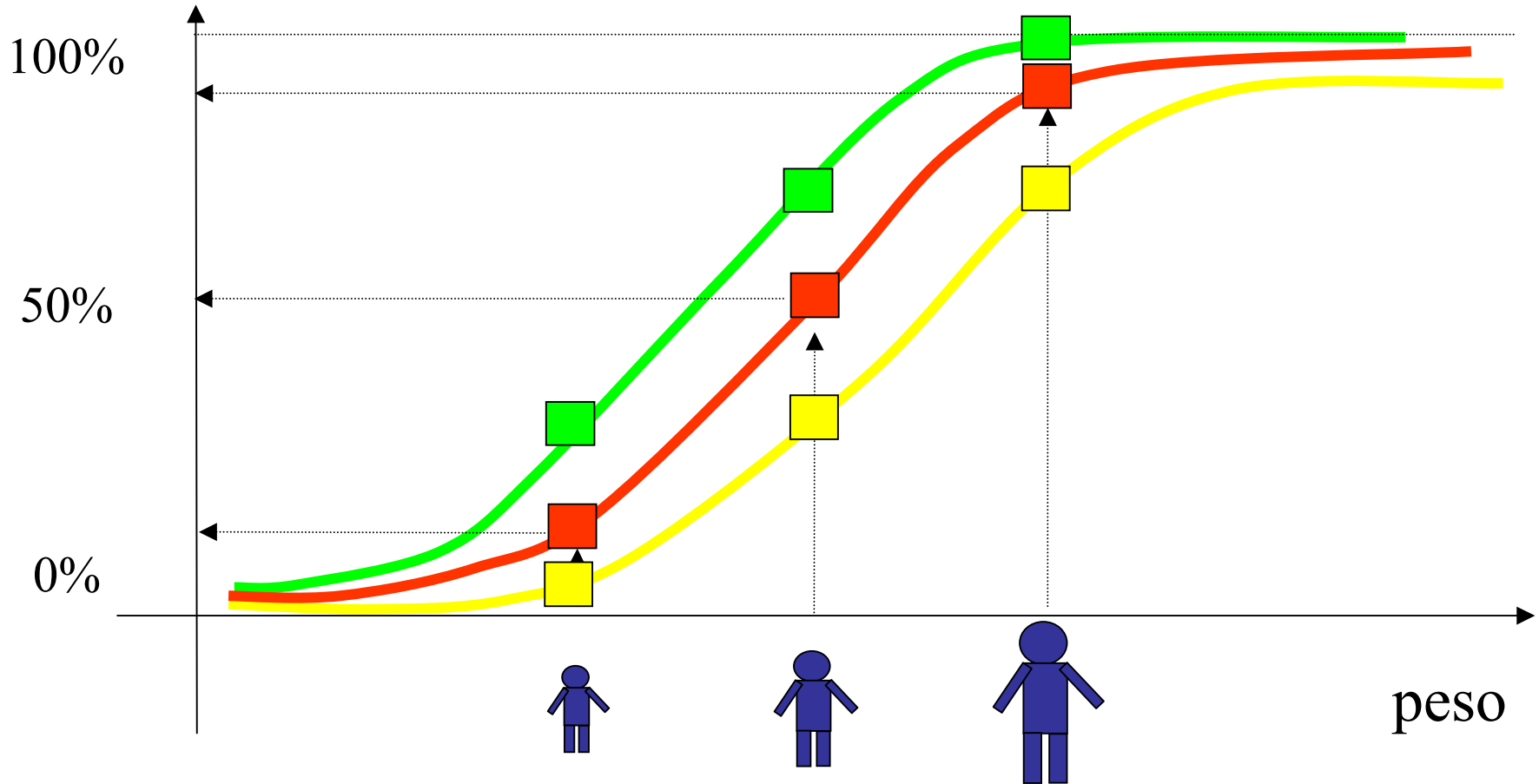


E se la pietra pesa meno?

Probabilità che il risultato sia = 1



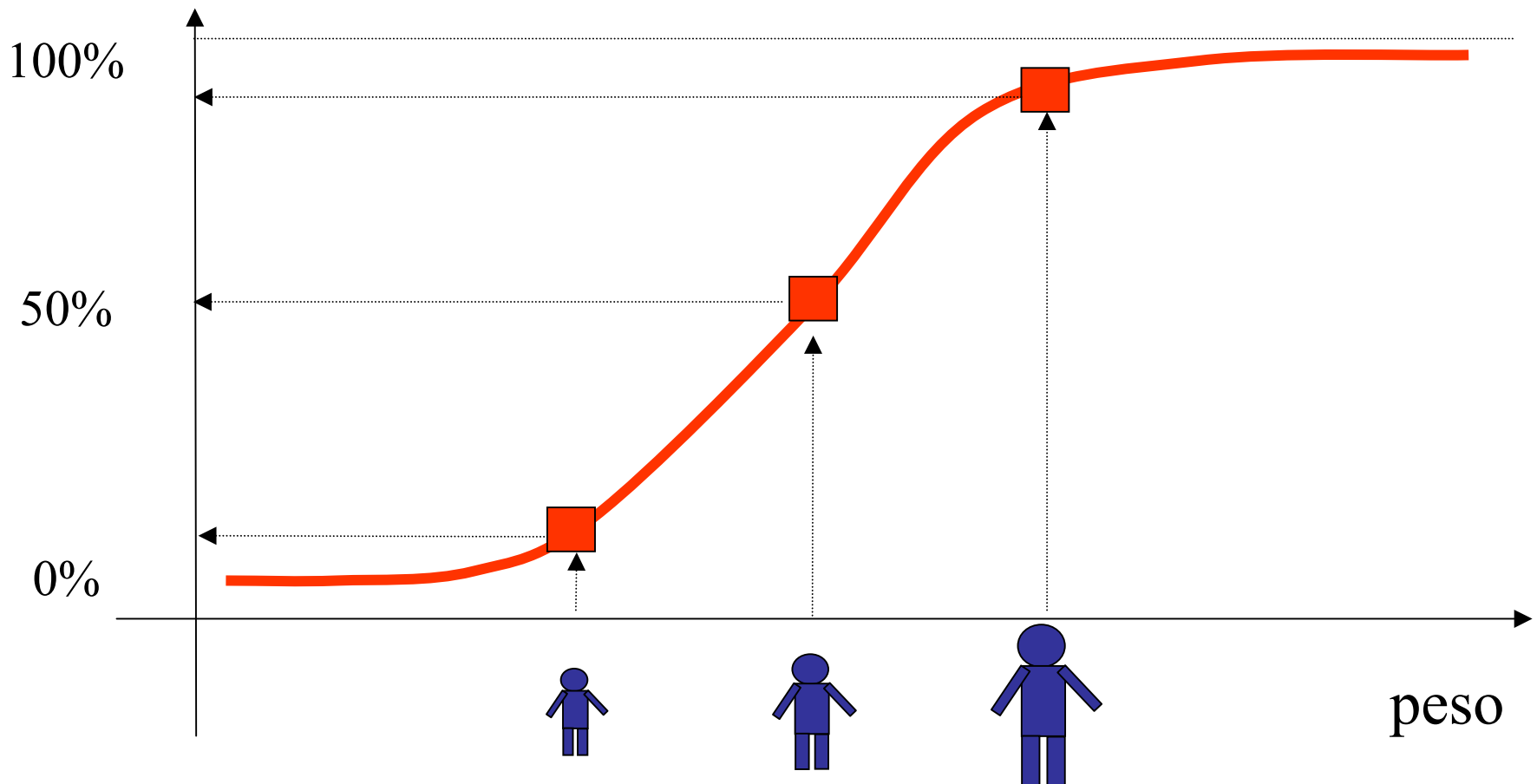
Joint Research Centre



Quali proprietà devono avere queste curve per rispettare la “*Specific Objectivity*” ?

Joint Research Centre

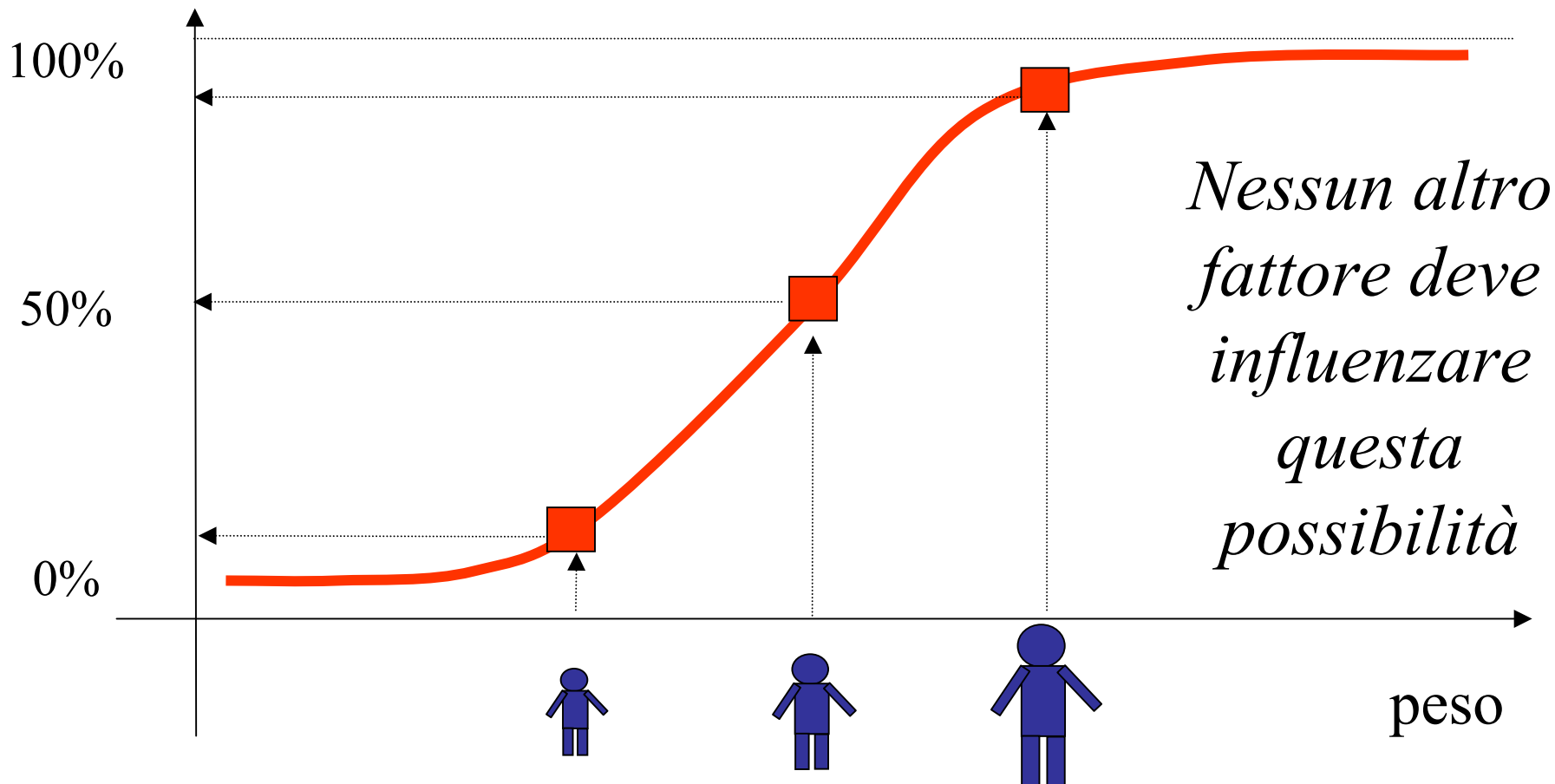
Probabilità che il risultato sia = 1

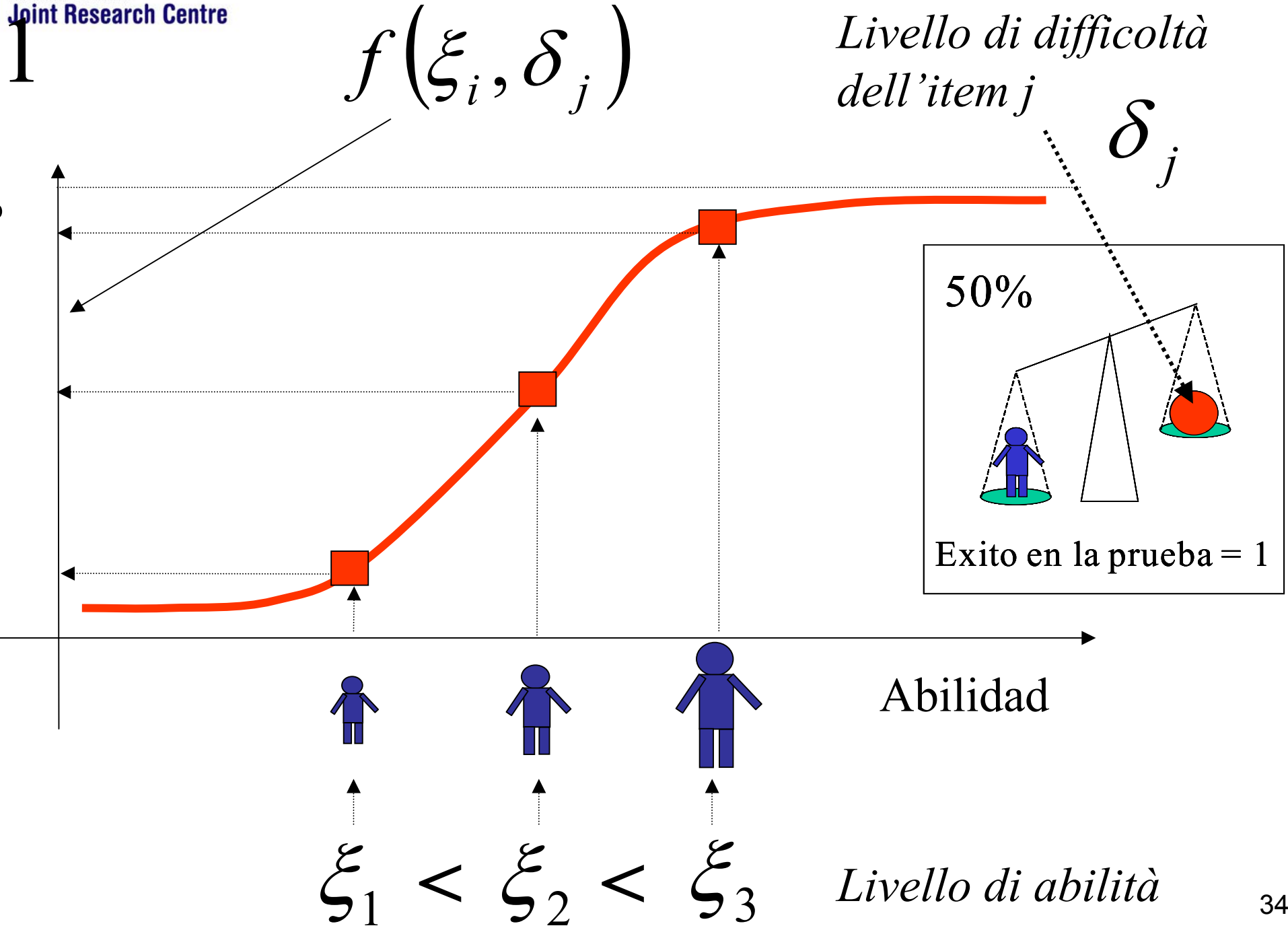


1 La probabilità di rispondere correttamente all'item deve dipendere solamente dalla **Abilità** dello studente e dalla **Difficoltà** dell'item

Joint Research Centre

Probabilità che il risultato sia = 1

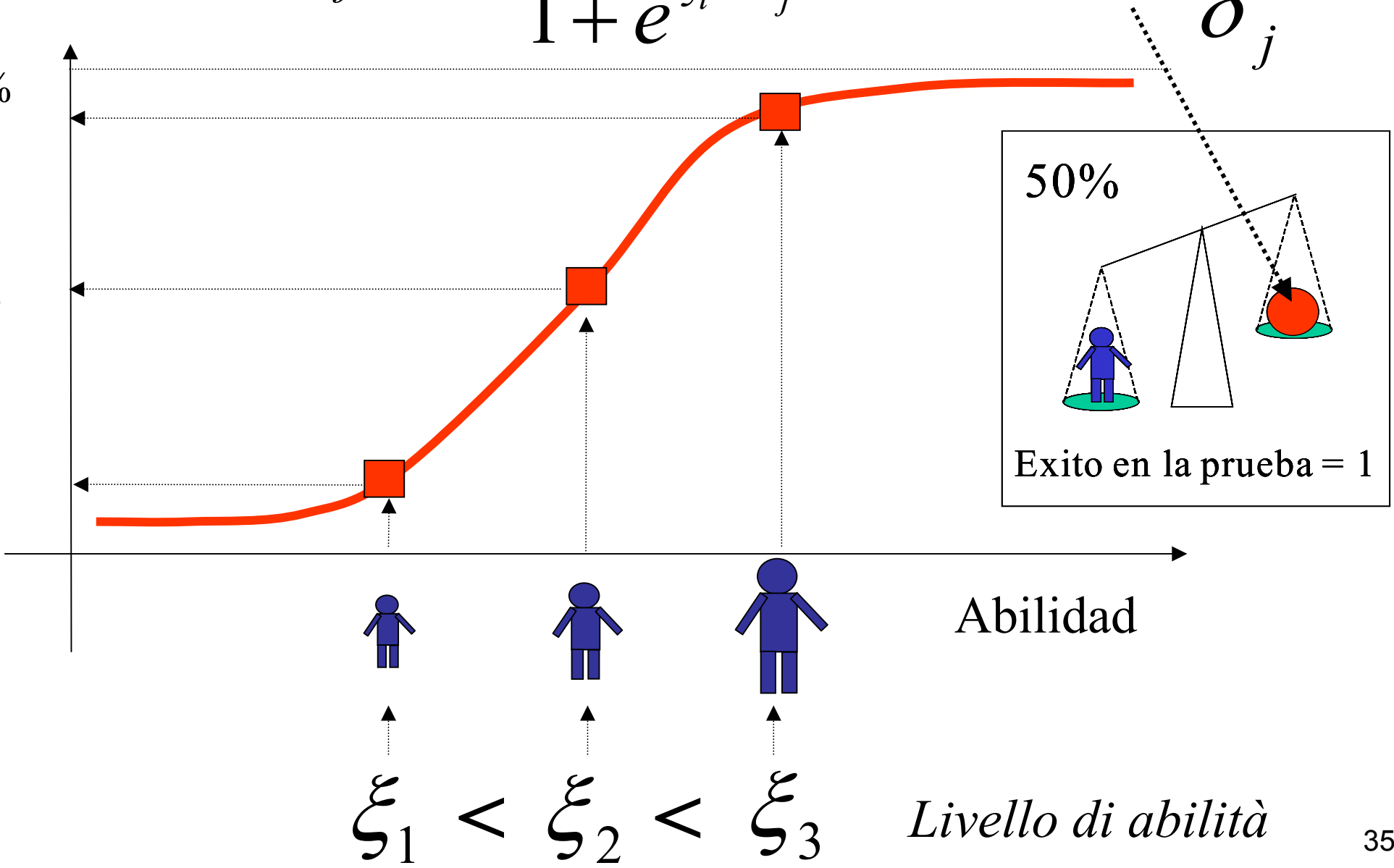




$$2 \quad f(\xi_i, \delta_j) = \frac{e^{\xi_i - \delta_j}}{1 + e^{\xi_i - \delta_j}}$$

Livello di difficoltà dell'item j

Joint Research Centre



Modelo de Rasch

- 1 ξ_i δ_j
- Abilità dello studente Difficoltà dell'item

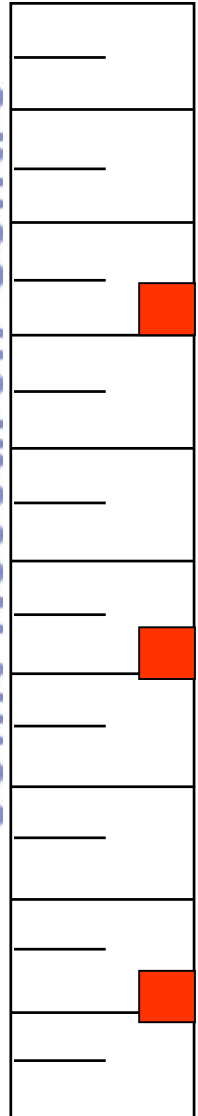
La probabilità che lo studente i risponda correttamente all'item j

2 $f(\xi_i, \delta_j) = \frac{e^{\xi_i - \delta_j}}{1 + e^{\xi_i - \delta_j}}$

Ciò significa che...

il “peso” stimato degli item
(difficoltà)

Joint Research Centre



$\frac{\partial y}{\partial x} = a \cdot (\ln x)^2 + b \cdot x, y = ?$ 3.21

$a \cdot x^2 + b \cdot x + c = ?$ 1.65

$2 + 2 = ?$ -3.57

È sempre
il
medesimo
nel tempo
e nello
spazio

Abilità dello
 studente

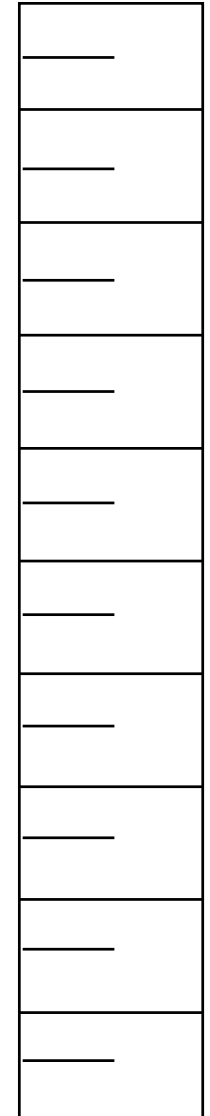
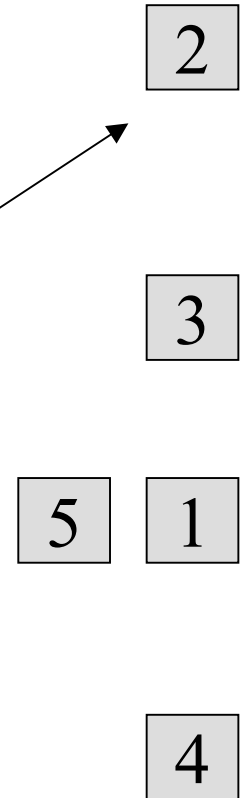
Scala RIT

Difficoltà
 dell'item

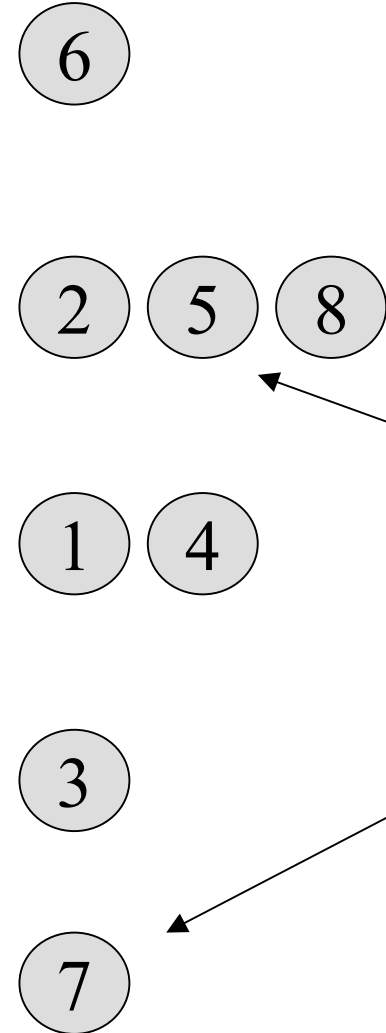
Molto abile

$\xi_1, \xi_2 \square \xi_n$

Poco abile



Difficile



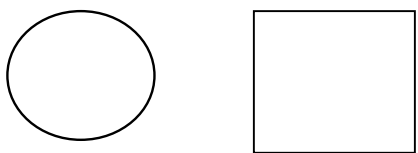
$\delta_1, \delta_2 \square \delta_k$

Facile

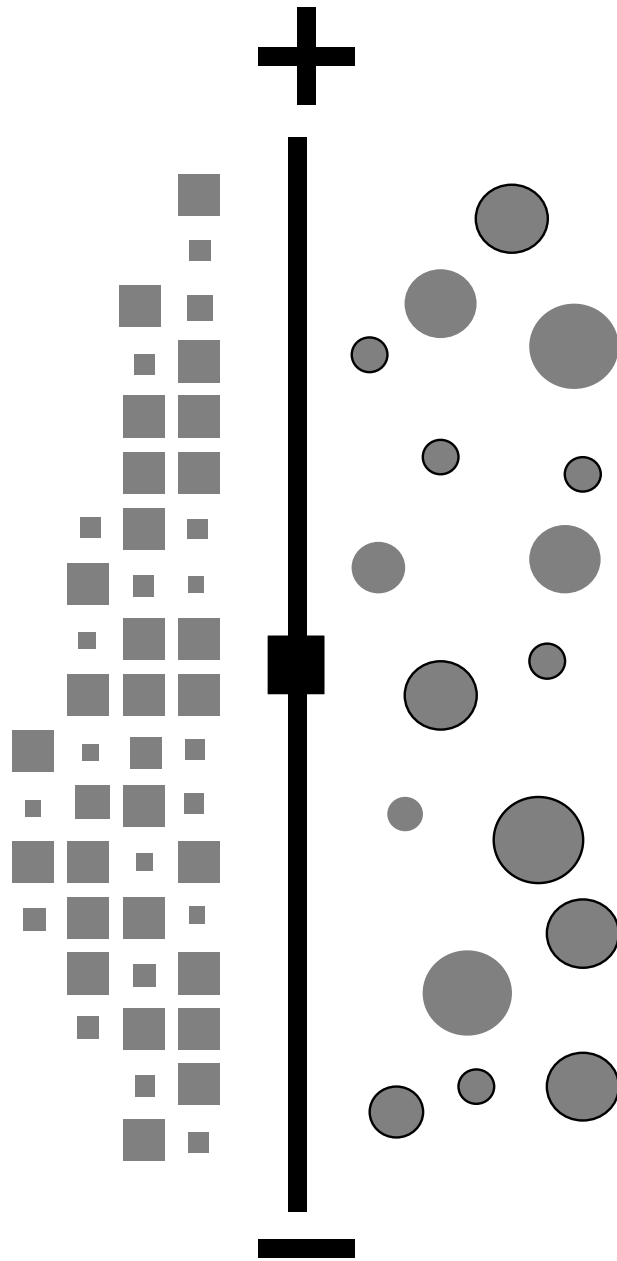
Risultati tipici dell'applicazione del modello di Rasch

Abilità dello
studente

Alto errore standard



Basso errore standard



difficoltà
dell'item



La traduzione in termini di criteri di valutazione diventa...

Australia

Joint Research Centre

ACER's Longitudi nal Literacy and Numeracy Study (LANNS)

Individual &
Normative
Progress

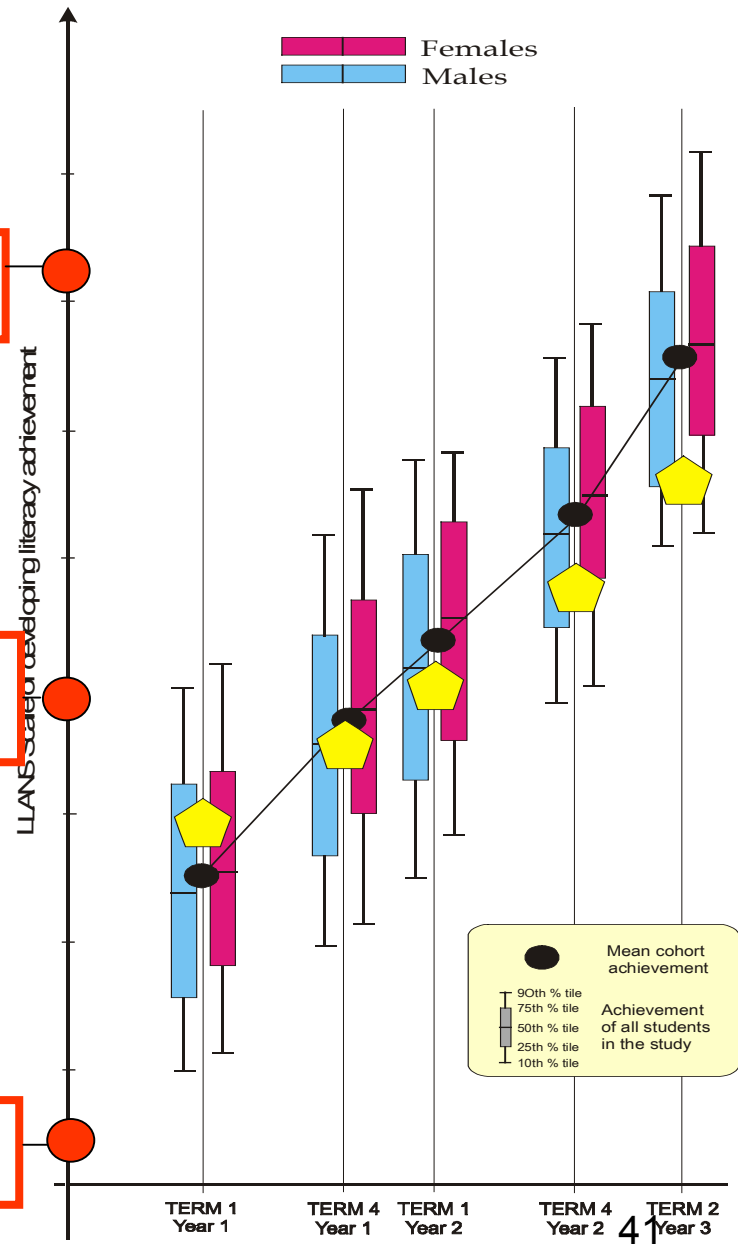
Longitudinal Literacy and Numeracy Study (LLANS)

LITERACY SCALE DESCRIPTION & NORMATIVE DISTRIBUTIONS

Angelico Jefferson 
 Warra School of Excellence

Note: The indicators listed on this side of the scale have been derived from the tasks completed in the LLANS assessments. Only a selected sample of these indicators has been used to describe developing achievement in literacy.

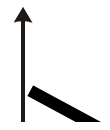
- Writes a variety of simple sentences; selects and controls content of own writing. Listens to a text and infers the reason for an event without picture clues. Uses full stops and capital letters to separate sentences. Identifies the purpose of parts of a text (eg, glossary, caption).
- Recognises implied meaning in a short section of a simple written text. Reads with word-for-word accuracy, an unseen, illustrated reader with a narrative structure, varied sentences and a wide range of common vocabulary. Segments and blends to pronounce unfamiliar words correctly. Spells some common words with irregular patterns, eg, 'Basket'. Controls content in writing, eg, selects specific details appropriate to the piece, or includes some explanations, opinions or reasons.
- Explains a story complication and resolution in a picture story book. Links images and text to construct meaning from own reading or listening. Reads with word-for-word accuracy, an unseen, factual early reader with a repetitive structure, varied content and some support from illustrations. Spells high frequency words with a range of patterns. Writes a piece that shows some overall coherence, eg, a sequence of events or a detailed list.
- From own reading or listening, identifies and explains key events, and follows steps in procedures of a picture story book and early readers. Reads common words with difficult spelling patterns, eg, 'because'. Spells some high frequency words with common patterns. Manipulates sounds in words, eg, swaps 'm' in 'smell' with 'p' to make 'spell'. Joins simple sentences using conjunctions.
- Offers simple explanations for a character's behavior, and locates explicit details from own reading or listening to picture story books. Reads unseen early readers with moderate accuracy (ie, omissions or substitutions do not consistently maintain meaning of text. Writes simple sentences that are mostly readable using phonetically plausible spelling for most common words. Lists ideas with little elaboration..
- After listening to a picture story book, includes several key aspects in a retelling. Reads simple common words. Identifies all the sounds in simple words. Writes one main idea with mostly recognisable words.
- Reads some very high frequency words. Recognises the same initial sounds in short words. Writes some recognisable words with spaces. Communicates some meaning in writing
- Describes an event or gives a limited retelling after listening to a picture story book. Reads a single word label by linking to the illustration. Names and sounds many letters. Writes own name correctly.
- Describes the main idea in an illustration after listening to a picture story book. Identifies writing and distinguishes words and letters. Writes a string of letters or scribble.
- Locates the front of a picture story book. Identifies a word.



Longitudinal Literacy and Numeracy Study (LLANS)

LITERACY SCALE DESCRIPTION & NORMATIVE DISTRIBUTIONS

Angelico Jefferson
Warra School of Excellence



Females
Males

Writes a variety of simple sentences; selects and controls content of own writing. Listens to a text and infers the reason for an event without picture clues. Uses full stops and capital letters to separate sentences. Identifies the purpose of parts of a text (eg, glossary, caption).

Literacy and

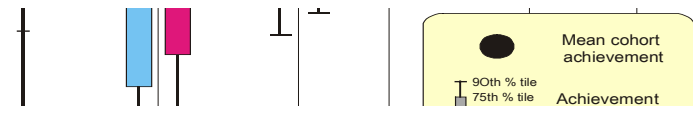
blends to pronounce unfamiliar words correctly. Spells some common words with irregular patterns, eg., 'Basket'. Controls content in writing, eg, selects specific details appropriate to the piece, or includes some explanations, opinions or reasons.



From own reading or listening, identifies and explains key events, and follows steps in procedures of a picture story book and early readers. Reads common words with difficult spelling patterns, eg, 'because'. Spells some high frequency words with common patterns. Manipulates sounds in words, eg, swaps 'm' in 'smell' with 'p' to make 'spell'. Joins simple sentences using conjunctions

Individual & Normative

all the sounds in simple words. Writes one main idea with mostly recognisable words. Reads some very high frequency words. Recognises the same initial sounds in short words. Writes some recognisable words with spaces. Communicates some meaning in writing



Describes the main idea in an illustration after listening to a picture story book. Identifies writing and distinguishes words and letters. Writes a string of letters or scribble.

Locates the front of a picture story book. Identifies a word.

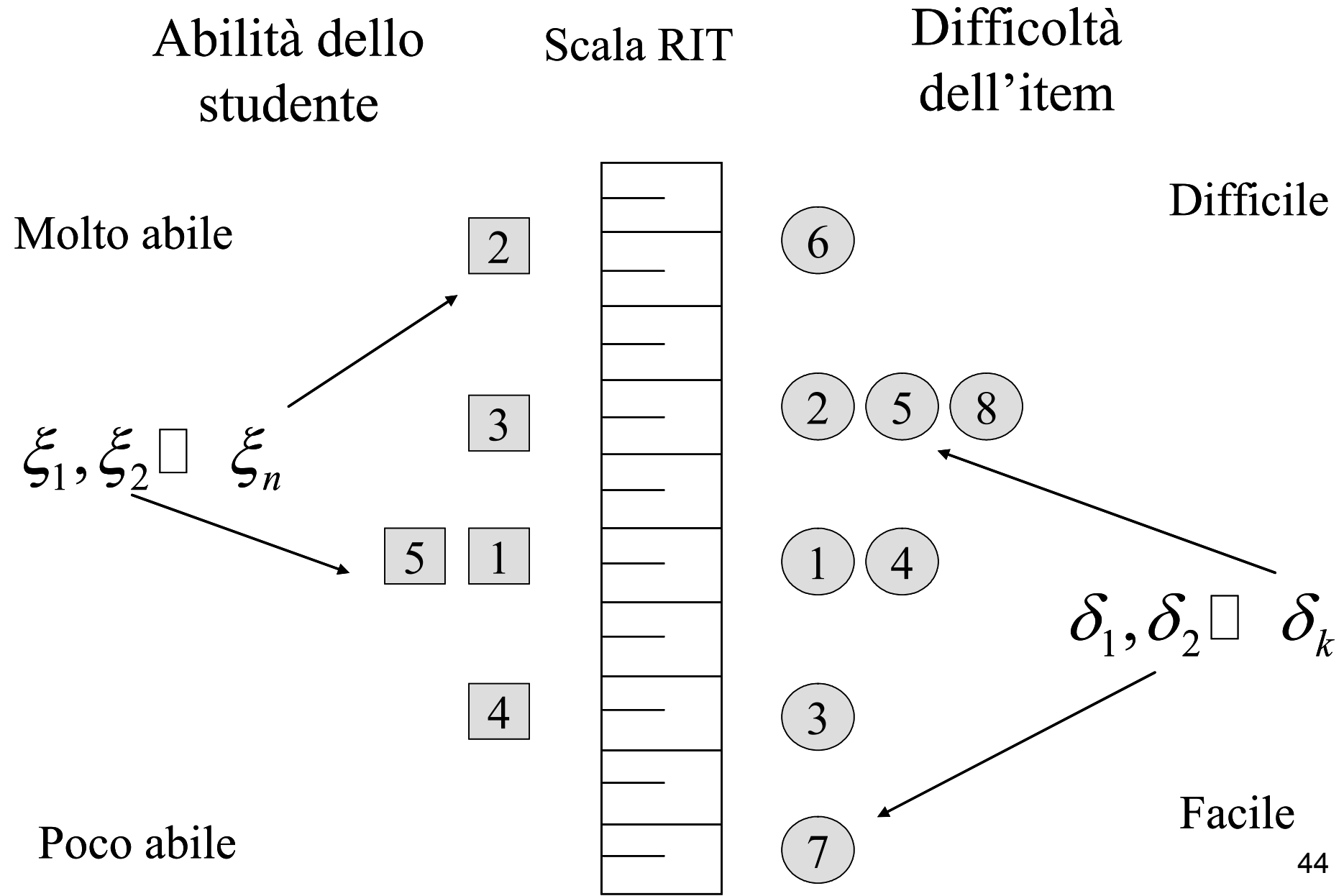


**Ma c'è un'ulteriore
possibilità...**

Il modello multifacet

Modello di Rasch tradizionale

Joint Research Centre



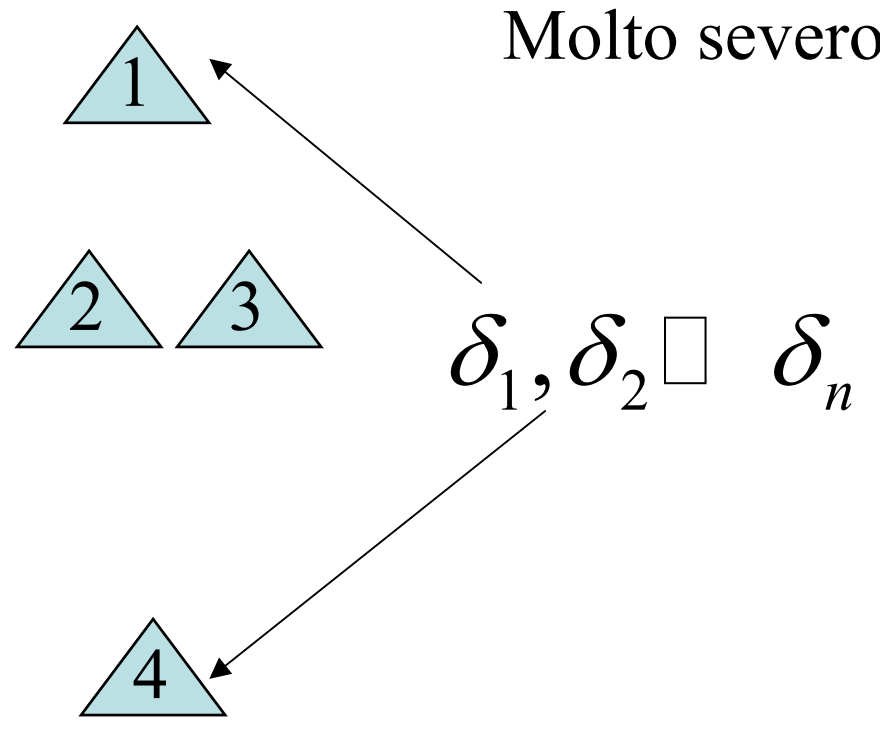
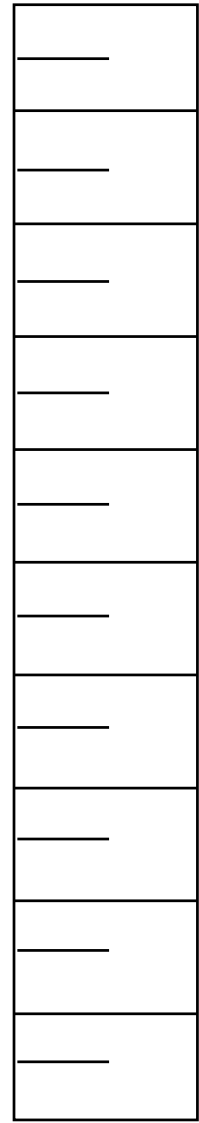
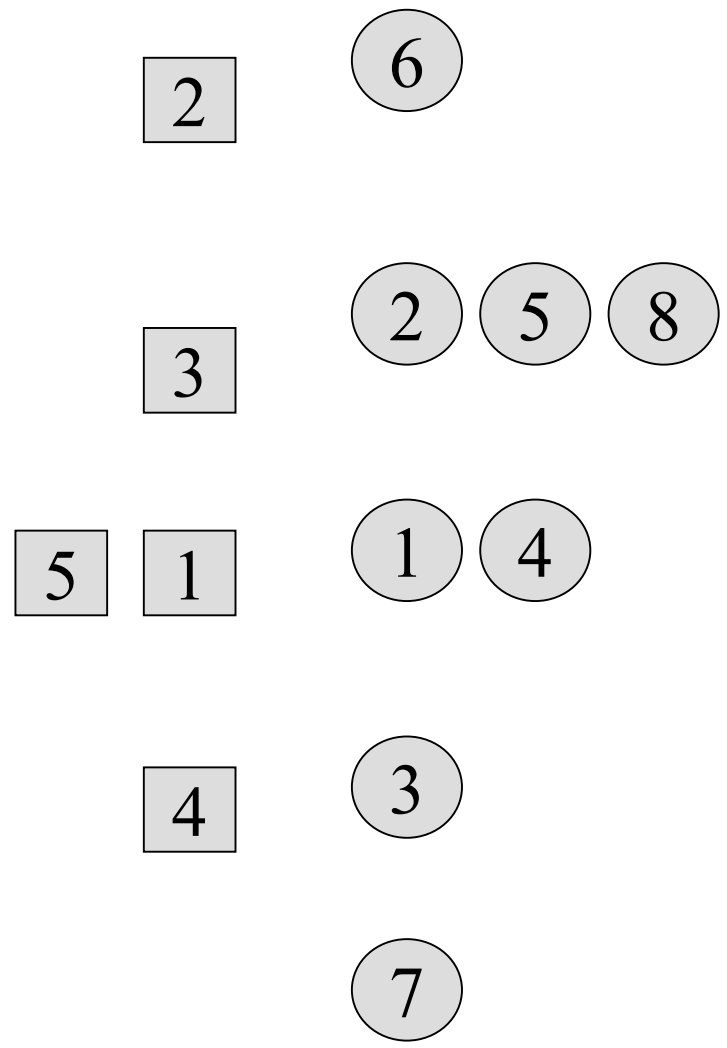
Scala RIT

Abilità studente

Difficoltà item

Severità giudice

Joint Research Centre





Per far questo, però, è necessario avere...

**DEFINIZIONE UNIVOCA DI
DIFFICOLTA'**

**E DI QUALITA' DELLA
RISPOSTA**



Il problema della valutazione delle performance

L'approccio corrente



Obiettivi formativi per il primo biennio della scuola primaria

Gli obiettivi formativi fino al primo biennio. Per questo, nel primo anno e nel primo biennio, vanno sempre esperiti a partire da problemi ed attività ricavati dall'esperienza diretta dei fanciulli. Tali problemi ed attività, per definizione, sono sempre unitarie e sintetiche, quindi mai riducibili né ad esercizi segmentati ed artificiali, né alla comprensione assicurata da singole prospettive disciplinari o da singole 'educazioni'. Richiedono, piuttosto, sempre, la mobilitazione di sensibilità e prospettive pluri, inter e transdisciplinari, nonché il continuo richiamo all'integralità educativa. Inoltre, aspetto ancora più importante, esigono che siano sempre dotate di senso, e quindi motivanti, per chi le svolge. Sarà, allo stesso tempo, preoccupazione dei docenti far scoprire agli allievi la progressiva possibilità di aggregare i quadri concettuali a mano a mano ricavati dall'esperienza all'interno di repertori via via più formali, che aprano all'ordinamento disciplinare e interdisciplinare del sapere.

- **Problemi ed attività derivate da esperienza diretta**
- **Richiedono sensibilità e prospettive pluri, inter e transdisciplinari**
- **Richiamo ad integralità educativa**



Obiettivi formativi per il secondo biennio della scuola primaria

Gli obiettivi formativi nel secondo biennio. Nel secondo biennio, quasi a conclusione di un itinerario formativo che ha portato i fanciulli a scoprire riflessivamente nella loro unitaria e complessa esperienza personale e socio-ambientale la funzionalità interpretativa, sistematicamente ordinatoria e, soprattutto, critica della semantica e della sintassi disciplinari, è possibile cominciare a coniugare senso globale dell'esperienza personale e rigore del singolo punto di vista disciplinare, organicità pluri, inter e transdisciplinare e svolgimento sistematico delle singole discipline, integralità dell'educazione e attenzione a singoli e peculiari aspetti di essa. Per questo, senza voler mai abbandonare l'aggancio globale all'esperienza e l'integralità di ogni processo educativo, è possibile organizzare le singole attività scolastiche per discipline e per una o più 'educazioni' appartenenti all'unica Convivenza civile.

Scissione di esperienza educativa in singole attività scolastiche organizzate per discipline



Il risultato

pratico?

Un problema di mantra...

Elementi per la valutazione nelle diverse fasi del percorso educativo

**I Biennio
Primaria**

**II Biennio
Primaria**

**Secondaria
I grado**

Documento di riferimento:

Circolare ministeriale

n. 85 1994

Storia	Collocare nel tempo fatti ed esperienze vissute e riconoscere rapporti di successione esistenti tra loro
	Distinguere e confrontare alcuni tipi di fonte storica orale e scritta
	Leggere ed interpretare le testimonianze del passato presenti sul territorio
	Rilevare il rapporto di contemporaneità tra azioni e situazioni

Storia	Individuare elementi di contemporaneità, di sviluppo nel tempo e di durata nei quadri storici di civiltà studiati
	Comprendere ed usare termini specifici del linguaggio disciplinare
	Collocare nello spazio gli eventi, individuando i possibili nessi tra eventi storici e caratteristiche geografiche di un territorio
	Scoprire radici storiche antiche della realtà locale
....	
....	

Storia	Utilizzare termini specifici del linguaggio disciplinare
	Comprendere aspetti essenziali della metodologia della ricerca storica e delle categorie di interpretazione storica
	Distinguere e selezionare vari tipi di fonte storica, ricavare informazioni da una o più fonti
	Scoprire specifiche radici storiche nella realtà locale e regionale

....	



Criteri per il I Biennio della scuola Primaria

Distinguere e confrontare alcuni tipi di fonte storica orale e scritta

Sf

Cosa significa?

Domanda: l'indipendenza americana

britanniche in terra nordamericana alla madrepatria, e si concluse con la costituzione da parte delle prime di una nazione indipendente, gli Stati Uniti d'America.

Le cause della guerra si possono identificare in una serie di imposte (Stamp Act, Townshend Acts) che volevano far gravare sulle colonie gli immensi costi di gestione dell'impero britannico.

L'approccio repressivo adottato dall'impero britannico nei confronti delle colonie (in particolare gli Intolerable Acts che imposero al Massachusetts un regime di occupazione militare) fecero stringere tutte le colonie americane attorno al Massachusetts e portarono, da una parte alla creazione di un esercito di 20000 uomini sotto il comando di Gorge Washington e dall'altra alla Dichiarazione d'Indipendenza.

La reazione della Gran Bretagna fu violenta ed il tentativo di Washington di bloccarla nelle isole di Long Island e Manhattan antistanti New York fu inizialmente vano ma Washington riuscì a riorganizzare le truppe ribelli ed infierire un primo grave colpo agli inglesi di stanza a Trenton nella notte di Natale del 1777 ed a recuperare l'intero New Jersey poco dopo. Inoltre, la collaborazione di Francia e Spagna (che dichiararono apertamente guerra all'Inghilterra) e di Olanda (che siglò una "pace amica" con i ribelli) implicarono un allargamento del fronte della guerra e posero l'Inghilterra in una posizione di svantaggio militare ed economico.

Il risultato fu una crisi economica in Inghilterra a conseguenza della quale l'esistente governo cadde. Il nuovo governo, creato da William Pitt, comprese che l'unica via per evitare il collasso era ridurre la pressione. Per questo vennero realizzate una serie di riforme economiche e, soprattutto, il 3 settembre del 1783 venne firmato il Trattato di Parigi con cui nacquero gli Stati Uniti



Il problema inespresso: L'univocità del giudizio

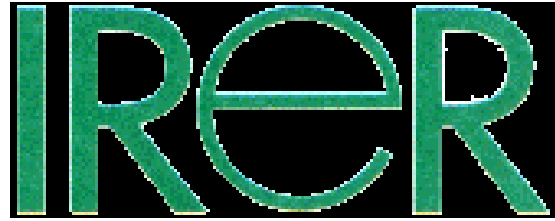
Ovvero...

La scissione dimensionale



Molteplicità di dimensioni

- 1. Indicazione sulla quantità e sulla qualità dei dati forniti**
- 2. Indicazione sulla correttezza delle conoscenze espresse in merito all'argomento**
- 3. Indicazione sulla consequenzialità storico-cronologica dell'esposizione**
- 4. Indicazione sul linguaggio**
- 5. Indicazione sulla consequenzialità argomentativa**



Dalla differenza, l'equità

Misurare gli apprendimenti
nella scuola dell'autonomia



Griglia di valutazione test di storia

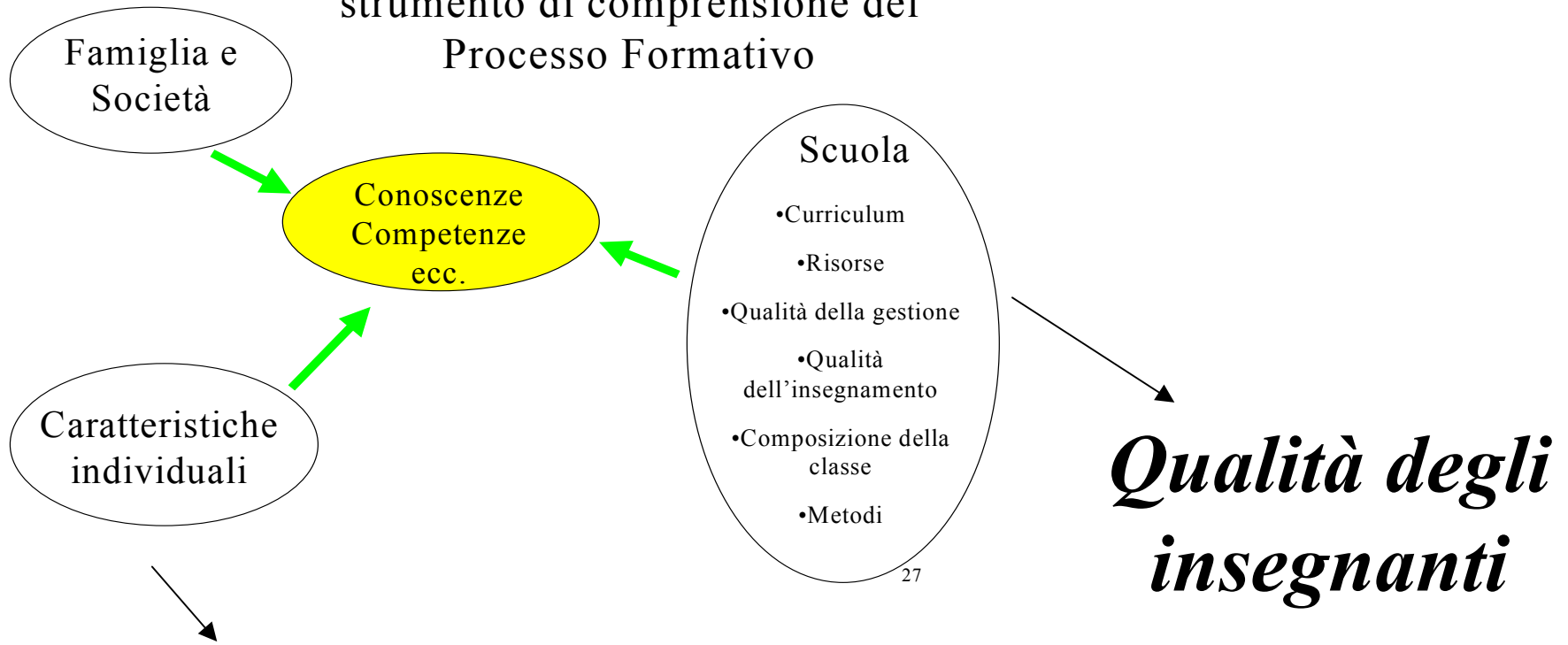
Cosa fa lo studente	Che tipo di intervento/giudizio è richiesto al correttore	Codifica (o scala) da usare
Sceglie e riporta il titolo di un argomento	narg – Segnare il numero dell'argomento prescelto	1,2, ...8 (ctr. le consegne della prova)
Esponde nel box eventi, nomi, date, luoghi ecc. che caratterizzano l'argomento	a – Indicazione sulla quantità e sulla qualità dei dati forniti	1= pochi dati e poco pertinenti 2= dati solo in parte pertinenti 3= dati pertinenti 4= dati ampi e pertinenti
Svolge in 20 righe l'argomento prescelto	b – Indicazione sulla correttezza delle conoscenze espresse in merito all'argomento	1=scorretta 2=imprecisa 3=fondamentalmente corretta 4=corretta e criticamente fondata
	c – Indicazione sulla consequenzialità storico-cronologica dell'esposizione	1=consequenzialità assente 2= consequenzialità presente ma errata 3=consequenzialità presente e corretta
	d – Indicazione sul linguaggio	1=scorrettezze (errori) 2=corretto 3=padronanza della sintassi e/o del lessico specifico
	e – Indicazione sulla consequenzialità argomentativa	1=confusa o assente 2=alcuni nessi imprecisi/mancanti 3=lineare 4=ben articolata



Fattori influenti: quelli veramente rilevanti

Valore Aggiunto:

strumento di comprensione del
Processo Formativo



27

Conoscenze pregresse dello studente



the
Mystery
of GOOD
by DAN GOLDHABER
Teaching

The evidence shows that good teachers make a clear difference in student achievement.

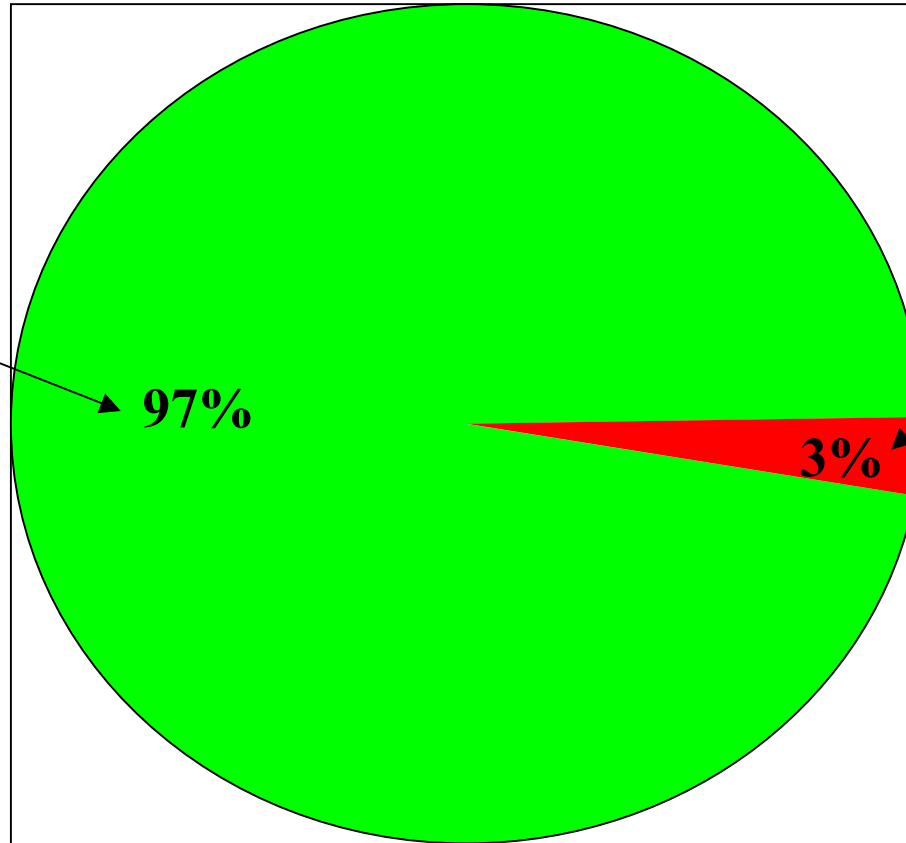
The problem is that we don't really know what makes A GOOD TEACHER



Fattori esplicativi dell'effetto dell'insegnante

**Caratteristiche
e
inosservabili:**

entusiasmo,
capacità di
trasmettere la
conoscenza
ecc.



**Caratteristiche
osservabili e
misurabili:**

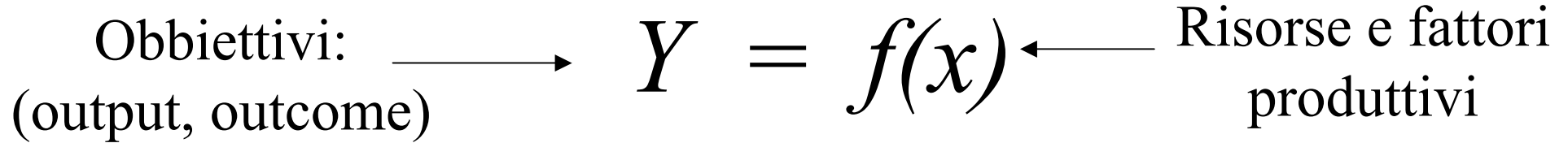
anni di
esperienza,
livello di
istruzione e
specializzazione,
performance
linguistiche

Accountability: efficienza ed efficacia

x Y	Fisico (o interno al sistema)	Monetario (o esterno al sistema)
Fisico	Efficacia interna Effetto del numero di studenti per classe o delle caratteristiche dell'insegnante sui risultati degli esami	Efficacia esterna In che misura certe pratiche pedagogiche influiscono sui salari dei diplomati?
Monetario	Efficienza interna A fini di accrescere i risultati è più conveniente ridurre il numero di studenti per classe o pagare di più gli insegnanti bravi?	Efficienza esterna E' più opportuno spendere risorse per l'istruzione universitaria o per la formazione professionale?

Accountability: efficienza ed efficacia

Capacità di raggiungere gli obiettivi a costi contenuti



Marlaine E. Lockheed & Eric Hanushek
 Concepts of Educational Efficiency and Effectiveness (1994)

Y	Fisico (o interno al sistema)	Monetario (o esterno al sistema)
x		
Fisico	Efficacia interna	Efficacia esterna
Monetario	Efficienza interna	Efficienza esterna

