EAM2025 XI Conference

23RD - 25TH JULY 2025

Spain Tenerife Canary Islands

European Association of Methodology

Constructing, improving, and shortening tests for skill assessment with competence-based test development

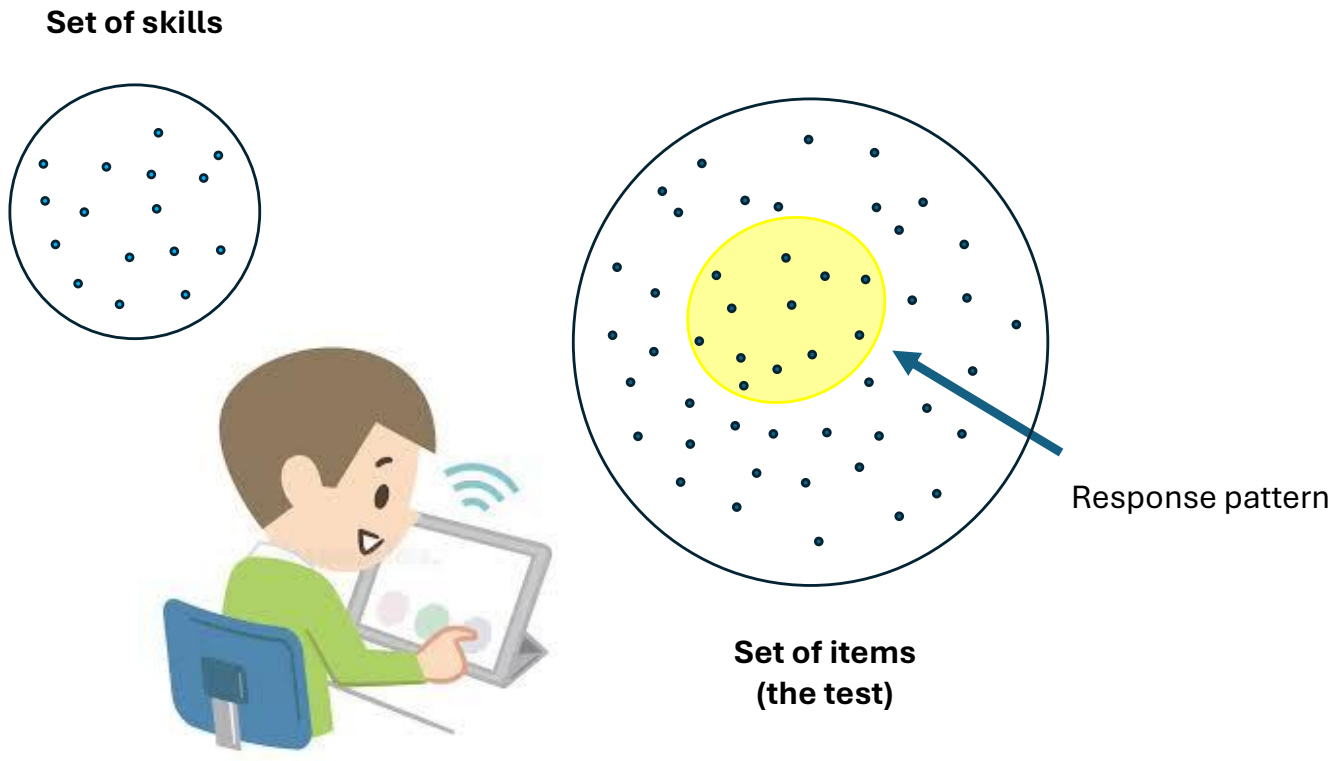
Pasquale Anselmi, Jürgen Heller, Luca Stefanutti, Egidio Robusto

Outline

- Competence-based knowledge space theory
- Competence-based test development
- Real-life applications of competence-based test development

Competence-based knowledge space theory (I)

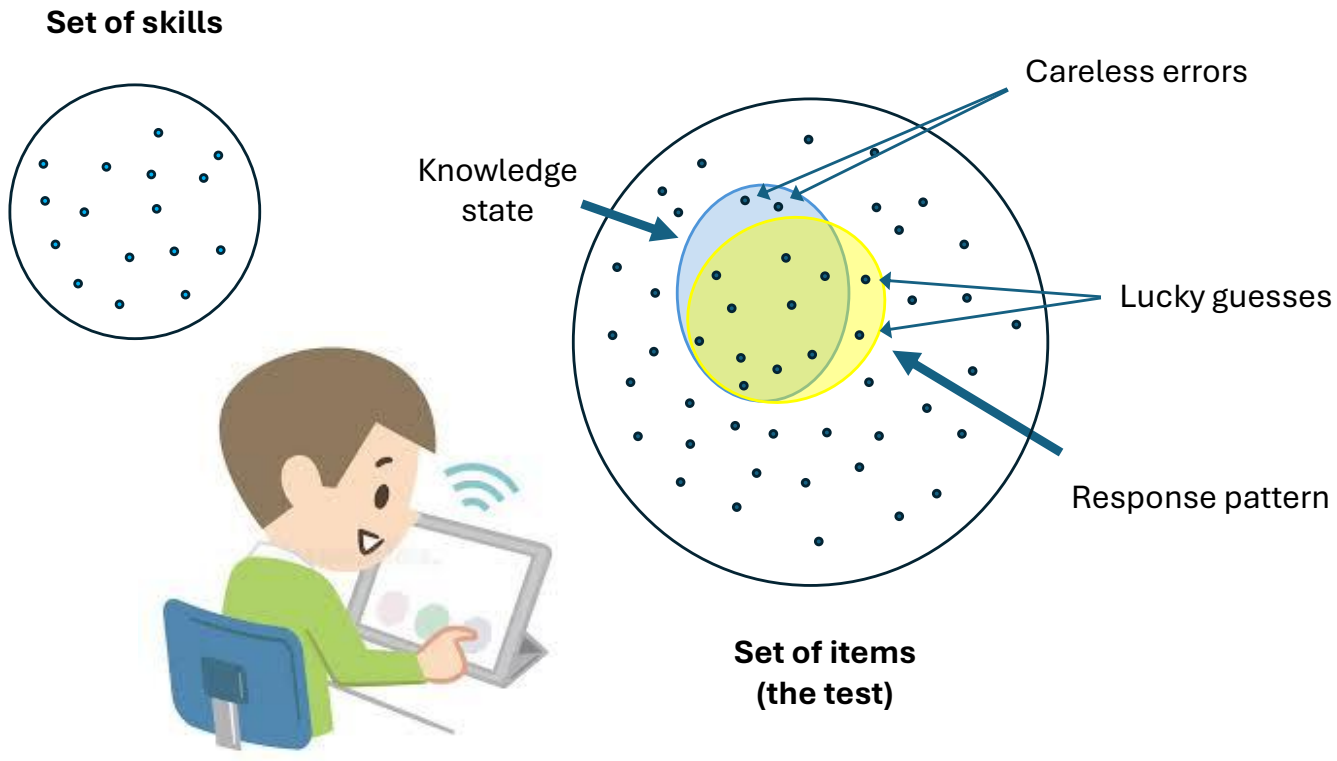
It allows for uncovering the set of skills an individual masters (the competence state) based on their response pattern



Competence-based knowledge space theory (I)

It allows for uncovering the set of skills an individual masters (the competence state) based on their response pattern

First, the set of items the individual is capable of solving (the knowledge state) is derived from the response pattern

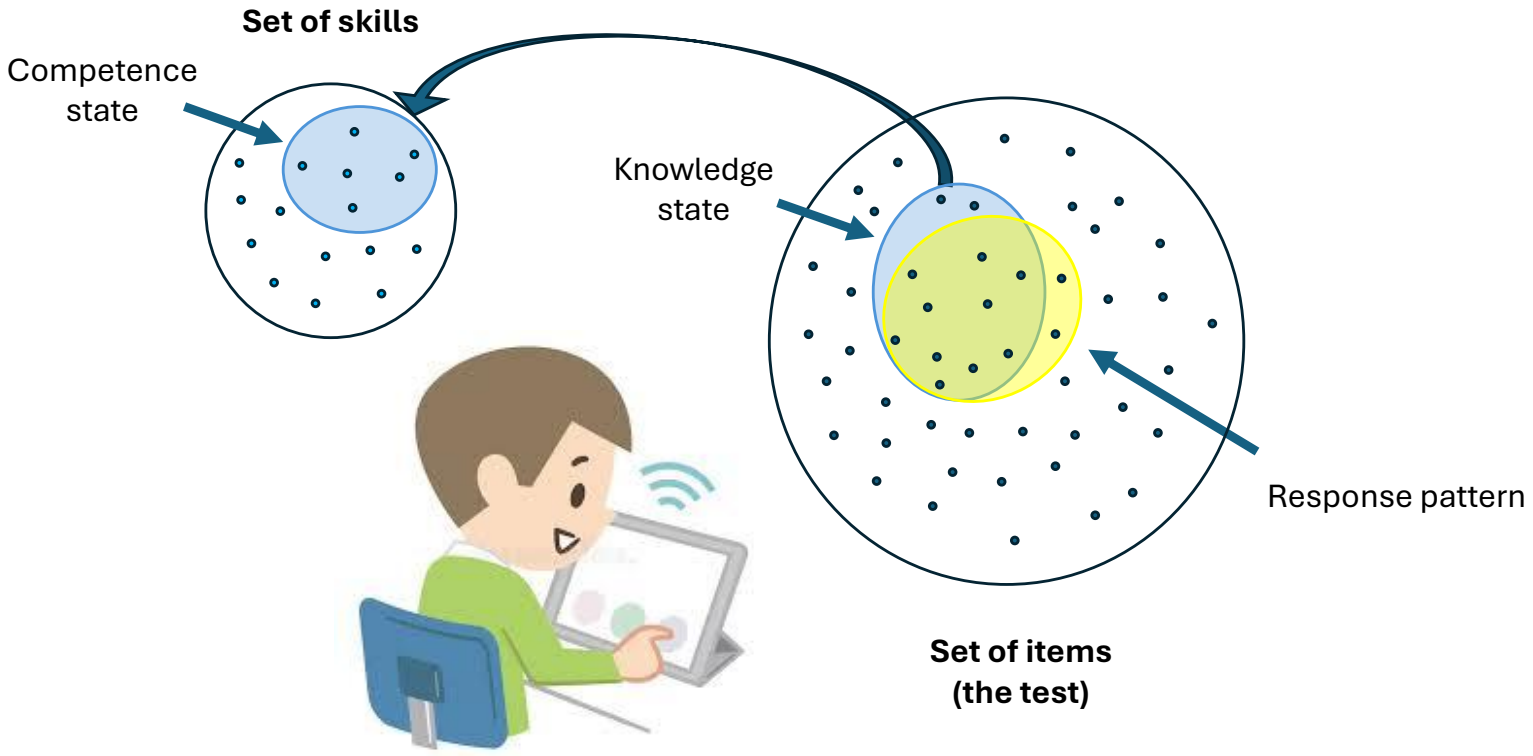


Competence-based knowledge space theory (I)

It allows for uncovering the set of skills an individual masters (the competence state) based on their response pattern

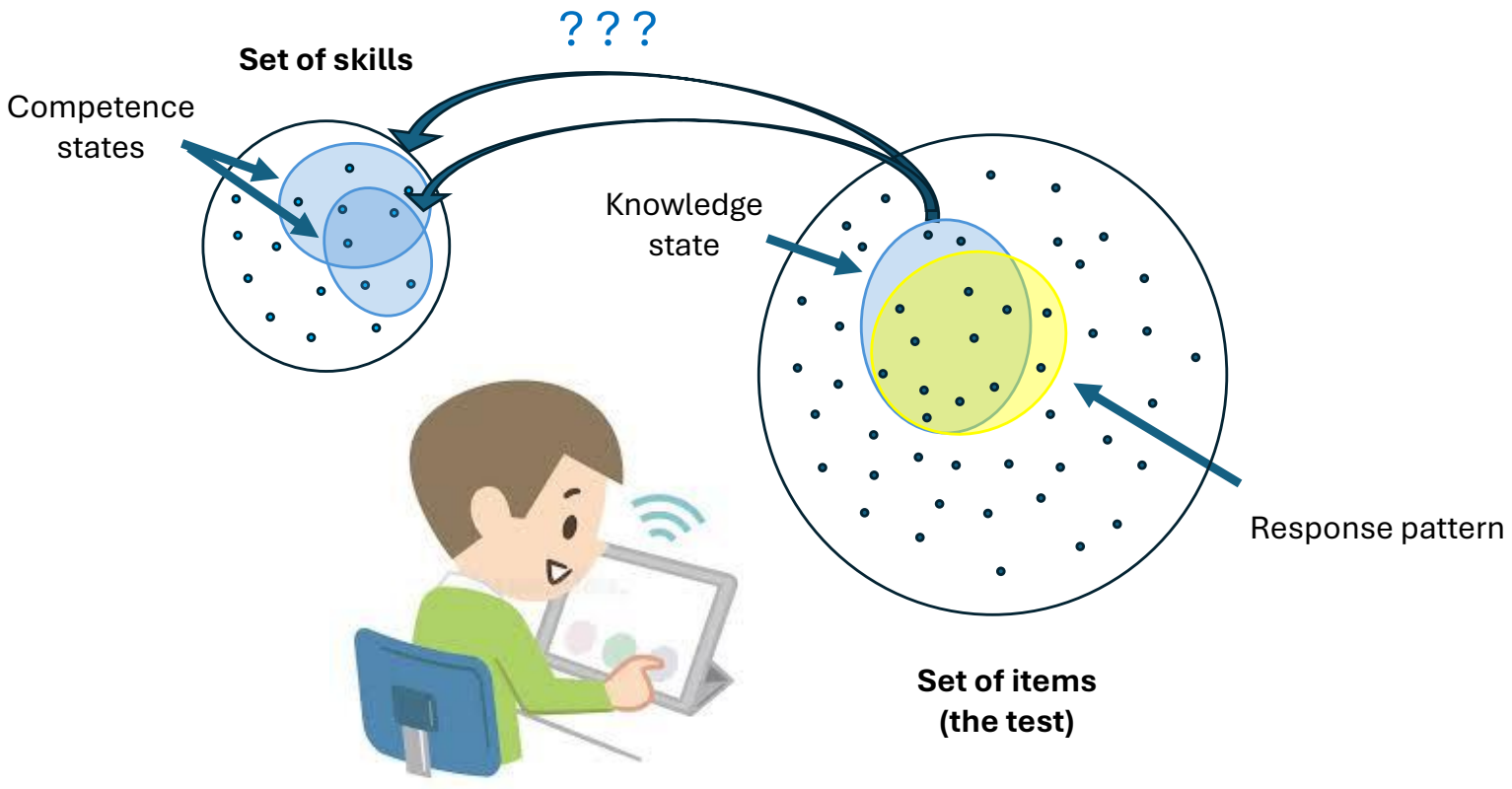
First, the set of items the individual is capable of solving (the knowledge state) is derived from the response pattern

Then, the competence state is inferred from the knowledge state



Competence-based knowledge space theory (II)

In some cases, more than one competence state may underlie a given knowledge state, meaning that the assessment of the skills an individual masters is not unique



Some basic notions on tests

A test is said to be:

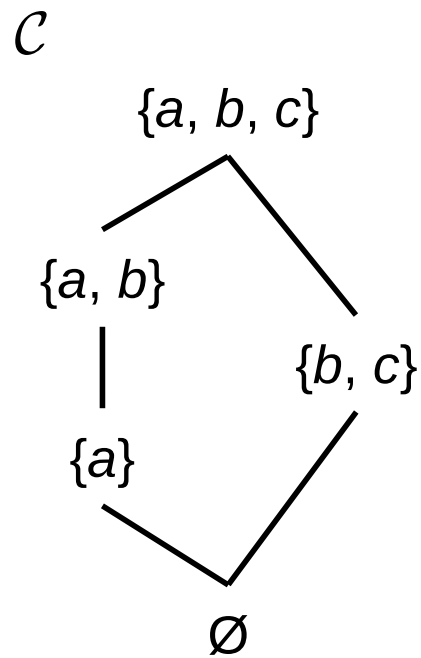
- conjunctive if all skills associated with an item – together forming its competency – are necessary to solve it. An individual can solve an item if its competency is a subset of their competence state
- fully informative if there are no two distinct competence states delineating the same knowledge state
- more informative than another test if it yields less uncertainty about the competence state given a knowledge state
- minimal if removing any item increases the number of competence states delineating the same knowledge state

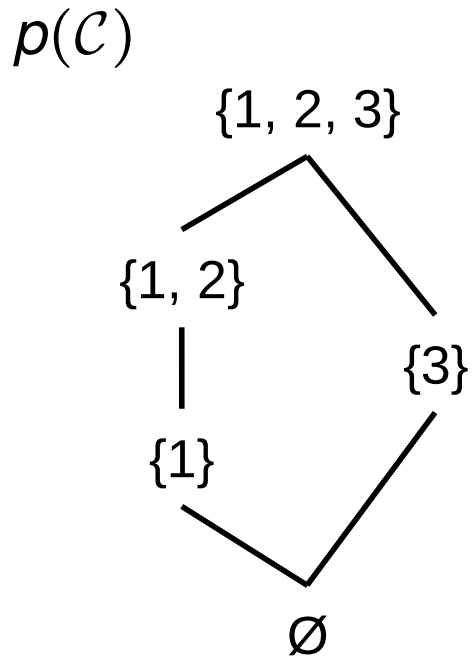
An example of a conjunctive test

Consider 5 competence states existing in a population

This test is fully informative ...

Item	Competency
1	$\{a\}$
2	$\{a, b\}$
3	$\{b, c\}$



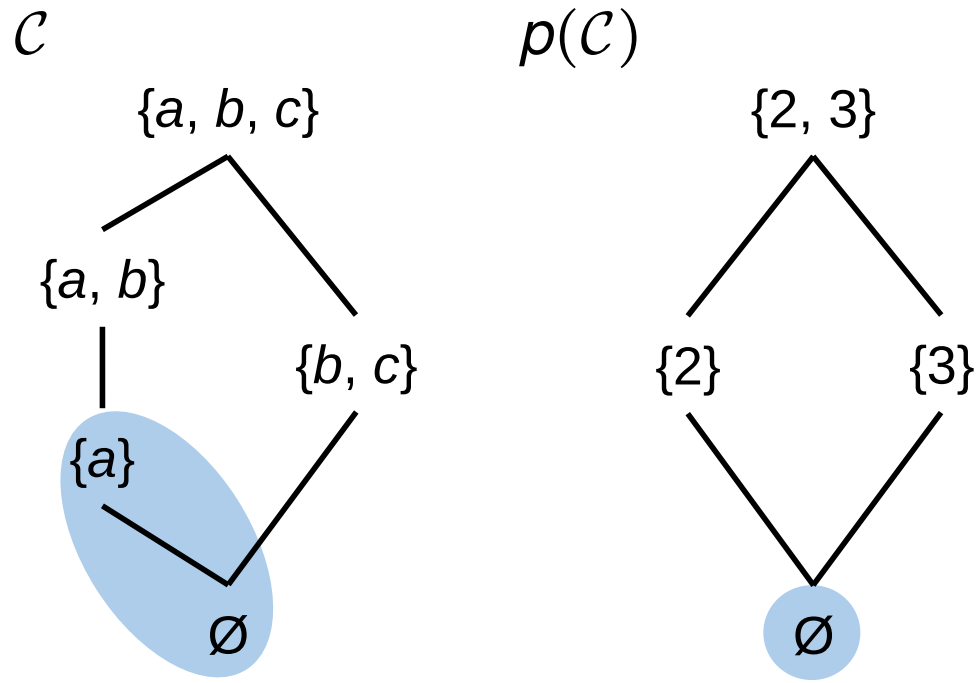


An example of a conjunctive test

Consider 5 competence states existing in a population

This test is fully informative and minimal

Item	Competency
1	$\{a\}$
2	$\{a, b\}$
3	$\{b, c\}$

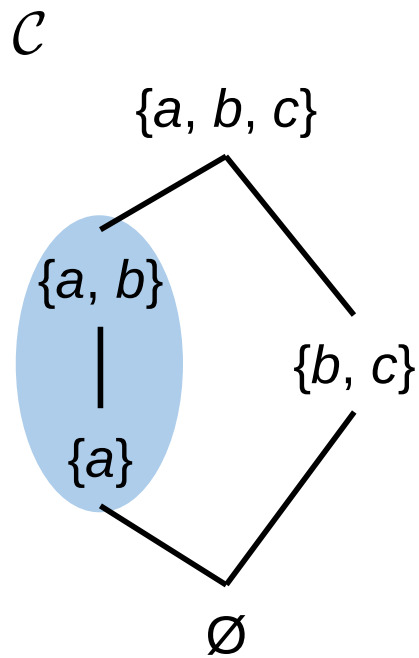


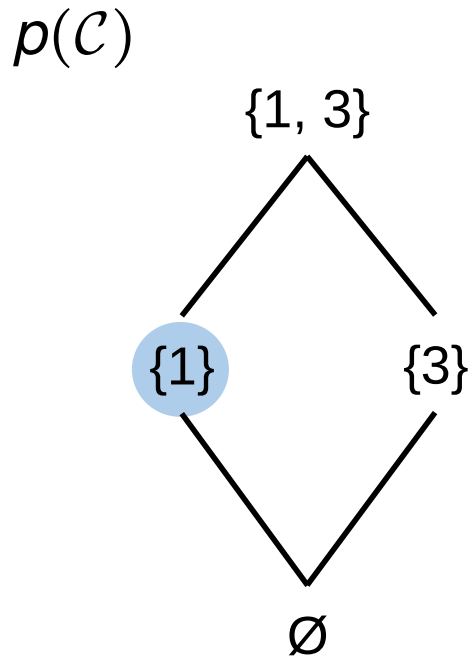
An example of a conjunctive test

Consider 5 competence states existing in a population

This test is fully informative and minimal

Item	Competency
1	$\{a\}$
2	$\{a, b\}$
3	$\{b, c\}$



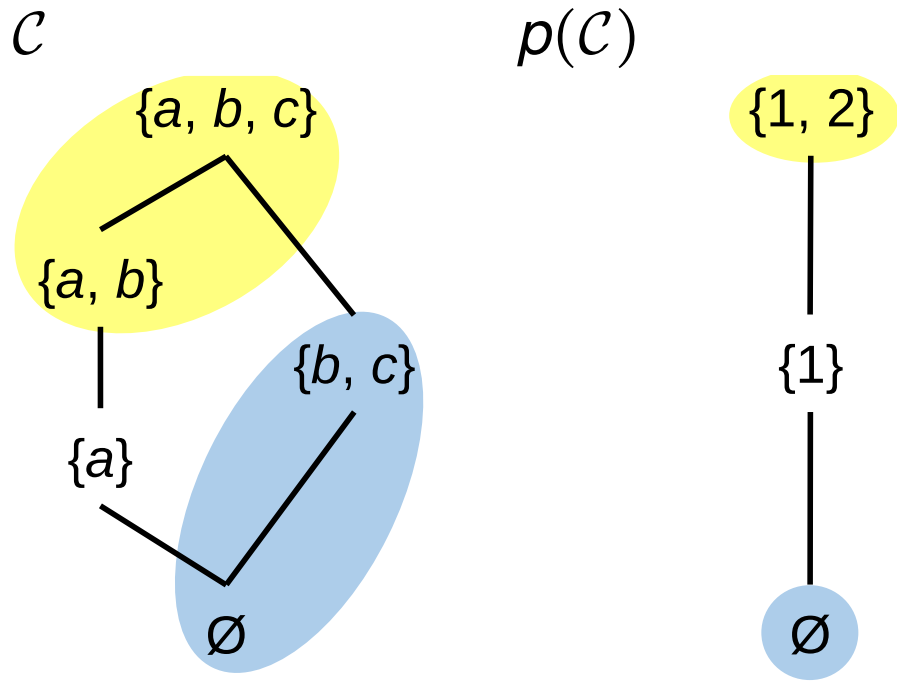


An example of a conjunctive test

Consider 5 competence states existing in a population

This test is fully informative and minimal

Item	Competency
1	$\{a\}$
2	$\{a, b\}$
3	$\{b, c\}$

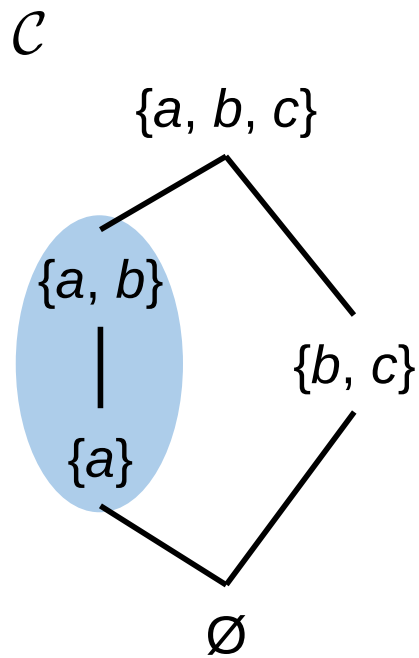


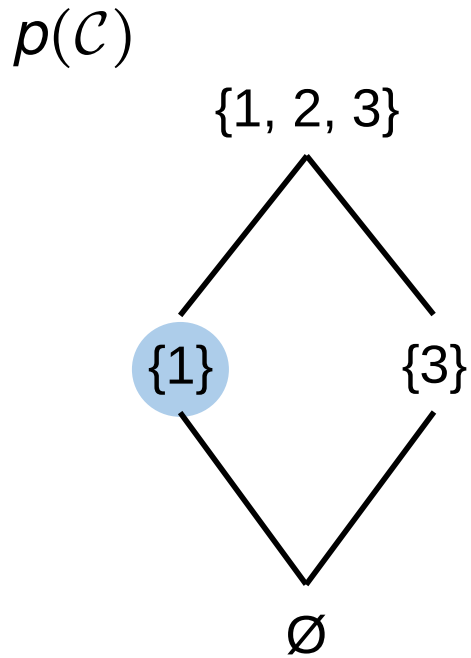
Another example of a conjunctive test

Consider 5 competence states existing in a population

This test is neither fully informative ...

Item	Competency
1	$\{a\}$
2	$\{a, b, c\}$
3	$\{b, c\}$





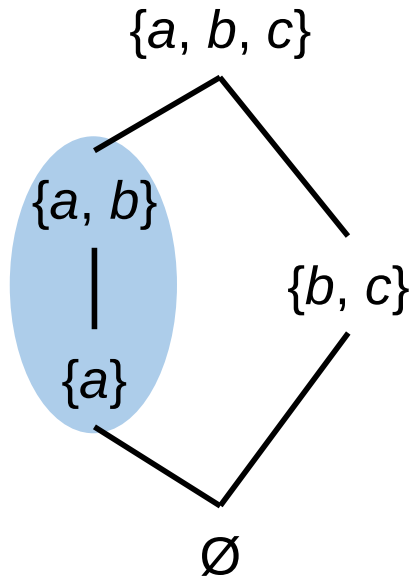
Another example of a conjunctive test

Consider 5 competence states existing in a population

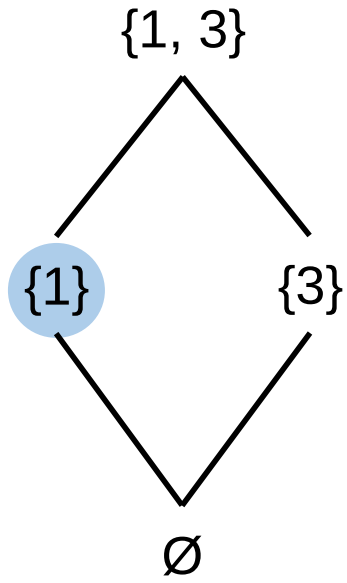
This test is neither fully informative nor minimal

Item	Competency
1	$\{a\}$
2	$\{a, b, c\}$
3	$\{b, c\}$

$\mathcal{C}$



$p(\mathcal{C})$



Competence-based test development

Competence-based test development (CbTD)[1] is a novel approach for constructing tests that are as informative as possible about individual's competence state

If desired, the tests can also be minimal

[1] Anselmi, Heller, Stefanutti, & Robusto (2022, 2024, 2025)

Discernibility function and reduct

Let:

- $\mathcal{D}$ be a nonempty set of pairs $\{C, C'\}$ of competence states
- $\mathcal{T}$ be a nonempty set of nonempty subsets T of skills, each interpreted as a competency

The discernibility function is a mapping δ that assigns to each pair $\{C, C'\} \in \mathcal{D}$ the set of all competencies $T \in \mathcal{T}$ that are a subset of one competence state but not the other. Thus, only one of C and C' enables solving an item requiring T

Every competency $T \in \delta(\{C, C'\})$ suffices to discern $\{C, C'\}$

A reduct is a minimal subset of $\mathcal{T}$ that discerns exactly the same pairs $\{C, C'\} \in \mathcal{D}$ as $\mathcal{T}$

An example of a reduct

Let:

- $\mathcal{D} = \{\{\emptyset, \{a, b\}\}, \{\{a, b\}, \{b, c\}\}, \{\{a, c\}, \{a, b, c\}\}\}$
- $\mathcal{T} = \{\{a\}, \{c\}, \{a, c\}, \{b, c\}, \{a, b, c\}\}$

The set $\mathcal{T}^{\diamond} = \{\{a\}, \{a, b, c\}\}$ is a reduct of $\mathcal{T}$

$\{C, C'\} \in \mathcal{D}$	$\delta(\{C, C'\})$	$\delta(\{C, C'\}) \cap \mathcal{T}^{\diamond}$
$\{\emptyset, \{a, b\}\}$	$\{\{a\}\}$	$\{\{a\}\}$
$\{\{a, b\}, \{b, c\}\}$	$\{\{a\}, \{c\}, \{b, c\}\}$	$\{\{a\}\}$
$\{\{a, c\}, \{a, b, c\}\}$	$\{\{b, c\}, \{a, b, c\}\}$	$\{\{a, b, c\}\}$

Real-life applications of CbTD

They cover the following scenarios:

- Constructing a test from scratch. Tests are obtained that are as informative as possible about individuals' competence states
- Improving an existing test. Lengthened forms of a test are obtained that are more informative about individuals' competence states than the original test
- Shortening an existing test. Shortened forms of a test are obtained that are as informative about individuals' competence states as the original test

Example of constructing a test from scratch (I)

Aim: Constructing a test for the assessment of students' mathematical formulation

Mathematical formulation is about structuring a sequence of appropriate symbols representing the quantitative relationships among the quantities present in a proposition expressed in verbal form

Example of constructing a test from scratch (II)

An expert[2] identified 8 skills relevant in mathematical formulation: (*a*) recognize useful and useless information, (*b*) compress multiple units into one text expression, (*c*) improve information ordering, (*d*) understand mathematical vocabulary and symbols, (*e*) translate a verbal expression into a numerical one and vice versa, (*f*) understand the characteristics of a relationship between quantities, (*g*) translate a verbal expression in a mathematical one when there are numbers and letters and vice versa, (*h*) recognize and explain relationships between variables

Moreover, the expert defined a prerequisite relationship among the 8 skills, which resulted in 26 competence states

[2] Giulia Barillari, high school math and physics teacher

Example of constructing a test from scratch (III)

A reduct consisting of 8 competencies was obtained

$\{a\}$
$\{a, c\}$
$\{d\}$
$\{b, d\}$
$\{b, d, e\}$
$\{b, d, g\}$
$\{h\}$
$\{f\}$

Example of constructing a test from scratch (III)

A reduct consisting of 8 competencies was obtained

Based on it, an 8-item test was constructed that is fully informative and minimal

Item	Competency	Item	Competency
1	$\{a\}$	5	$\{b, d, e\}$
2	$\{a, c\}$	6	$\{b, d, g\}$
3	$\{d\}$	7	$\{h\}$
4	$\{b, d\}$	8	$\{f\}$

Tested on the responses by 175 middle and high school students, the test showed good fit: $\chi^2(214) = 61.67, p = .48$

Example of improving an existing test (I)

Aim: Improving Booklet 2 of the TIMSS 2003 eight grade mathematics test by adding a minimal set of items appropriately selected from Booklet 1

Example of improving an existing test (II)

Booklet 2 consists of 30 items that are assumed to require 15 skills to be solved[3]

Item ID	Item	Competency	Item ID	Item	Competency	Item ID	Item	Competency
M012016	1	$\{d, i\}$	M022139	13	$\{a, b, h\}$	M032381	25	$\{c, g\}$
M012017	2	$\{f, k, m\}$	M022144	14	$\{m\}$	M032416	26	$\{c, n\}$
M022251	3	$\{e\}$	M022253	15	$\{e, f\}$	M032160	27	$\{b, g\}$
M022185	4	$\{e\}$	M022156	16	$\{b, f\}$	M032540	28	$\{f\}$
M022191	5	$\{f\}$	M022104	17	$\{d\}$	M032698	29	$\{e, f\}$
M022194	6	$\{b, m\}$	M022106	18	$\{a, f, g\}$	M032529	30	$\{b, h\}$
M022196	7	$\{e\}$	M022110	19	$\{c\}$			
M022198	8	$\{d\}$	M032307	20	$\{i, j, n\}$			
M022199	9	$\{c, i, l, n\}$	M032523	21	$\{j, n\}$			
M012025	10	$\{g\}$	M032701	22	$\{b, f\}$			
M012027	11	$\{a, b, h\}$	M032704	23	$\{a, b, g\}$			
M012029	12	$\{m\}$	M032525	24	$\{f, i\}$			

The booklet is not fully informative: The 860 competence states delineate only 329 knowledge states

[3] Su et al. (2013)

Example of improving an existing test (III)

By adding 5 items appropriately selected from Booklet 1, an improved version of Booklet 2 is obtained that is more informative than the original one: The 860 competence states delineate 640 knowledge states against the 329 originally obtained

Item ID	Item	Competency	Item ID	Item	Competency	Item ID	Item	Competency
M012016	1	$\{d, i\}$	M022139	13	$\{a, b, h\}$	M032381	25	$\{c, g\}$
M012017	2	$\{f, k, m\}$	M022144	14	$\{m\}$	M032416	26	$\{c, n\}$
M022251	3	$\{e\}$	M022253	15	$\{e, f\}$	M032160	27	$\{b, g\}$
M022185	4	$\{e\}$	M022156	16	$\{b, f\}$	M032540	28	$\{f\}$
M022191	5	$\{f\}$	M022104	17	$\{d\}$	M032698	29	$\{e, f\}$
M022194	6	$\{b, m\}$	M022106	18	$\{a, f, g\}$	M032529	30	$\{b, h\}$
M022196	7	$\{e\}$	M022110	19	$\{c\}$	M022043	31	$\{l\}$
M022198	8	$\{d\}$	M032307	20	$\{i, j, n\}$	M012042	32	$\{e, i\}$
M022199	9	$\{c, i, l, n\}$	M032523	21	$\{j, n\}$	M012001	33	$\{a, l, n\}$
M012025	10	$\{g\}$	M032701	22	$\{b, f\}$	M012041	34	$\{a\}$
M012027	11	$\{a, b, h\}$	M032704	23	$\{a, b, g\}$	M032570	35	$\{b\}$
M012029	12	$\{m\}$	M032525	24	$\{f, i\}$			

No further improvement of Booklet 2 is possible by exploiting other items in Booklet 1

Example of shortening an existing test (I)

Aim: Developing a shortened form of the fraction subtraction test[4] that is as informative about the competence states underlying the knowledge states as the full test

[4] Tatsuoka (1990)

Example of shortening an existing test (II)

The following 20 mixed-number subtraction items are assumed to require 8 skills to be solved[5]: (a) convert a whole number to a fraction, (b) separate a whole number from a fraction, (c) simplify before subtracting, (d) find a common denominator, (e) borrow from whole number part, (f) column borrow to subtract the second numerator from the first, (g) subtract numerators, (h) reduce answers to simplest form

Item		Competency	Item		Competency
1	$\frac{5}{3} - \frac{3}{4}$	$\{d, f, g\}$	11	$4\frac{1}{3} - 2\frac{4}{3}$	$\{b, e, g\}$
2	$\frac{3}{4} - \frac{3}{8}$	$\{d, g\}$	12	$1\frac{1}{8} - \frac{1}{8}$	$\{g, h\}$
3	$\frac{5}{6} - \frac{1}{9}$	$\{d, g\}$	13	$3\frac{3}{8} - 2\frac{5}{6}$	$\{b, d, e, g\}$
4	$3\frac{1}{2} - 2\frac{3}{2}$	$\{b, c, e, g\}$	14	$3\frac{4}{5} - 3\frac{2}{5}$	$\{b, g\}$
5	$4\frac{3}{5} - 3\frac{4}{10}$	$\{b, d, g, h\}$	15	$2 - \frac{1}{3}$	$\{a, g\}$
6	$\frac{6}{7} - \frac{4}{7}$	$\{g\}$	16	$4\frac{5}{7} - 1\frac{4}{7}$	$\{b, g\}$
7	$3 - 2\frac{1}{5}$	$\{a, b, g\}$	17	$7\frac{3}{5} - \frac{4}{5}$	$\{b, e, g\}$
8	$\frac{2}{3} - \frac{2}{3}$	$\{g\}$	18	$4\frac{1}{10} - 2\frac{8}{10}$	$\{b, e, f, g\}$
9	$3\frac{7}{8} - 2$	$\{b\}$	19	$4 - 1\frac{4}{3}$	$\{a, b, c, e, g\}$
10	$4\frac{4}{12} - 2\frac{7}{12}$	$\{b, e, g, h\}$	20	$4\frac{1}{3} - 1\frac{5}{3}$	$\{b, c, e, g\}$

[5] DeCarlo (2011)

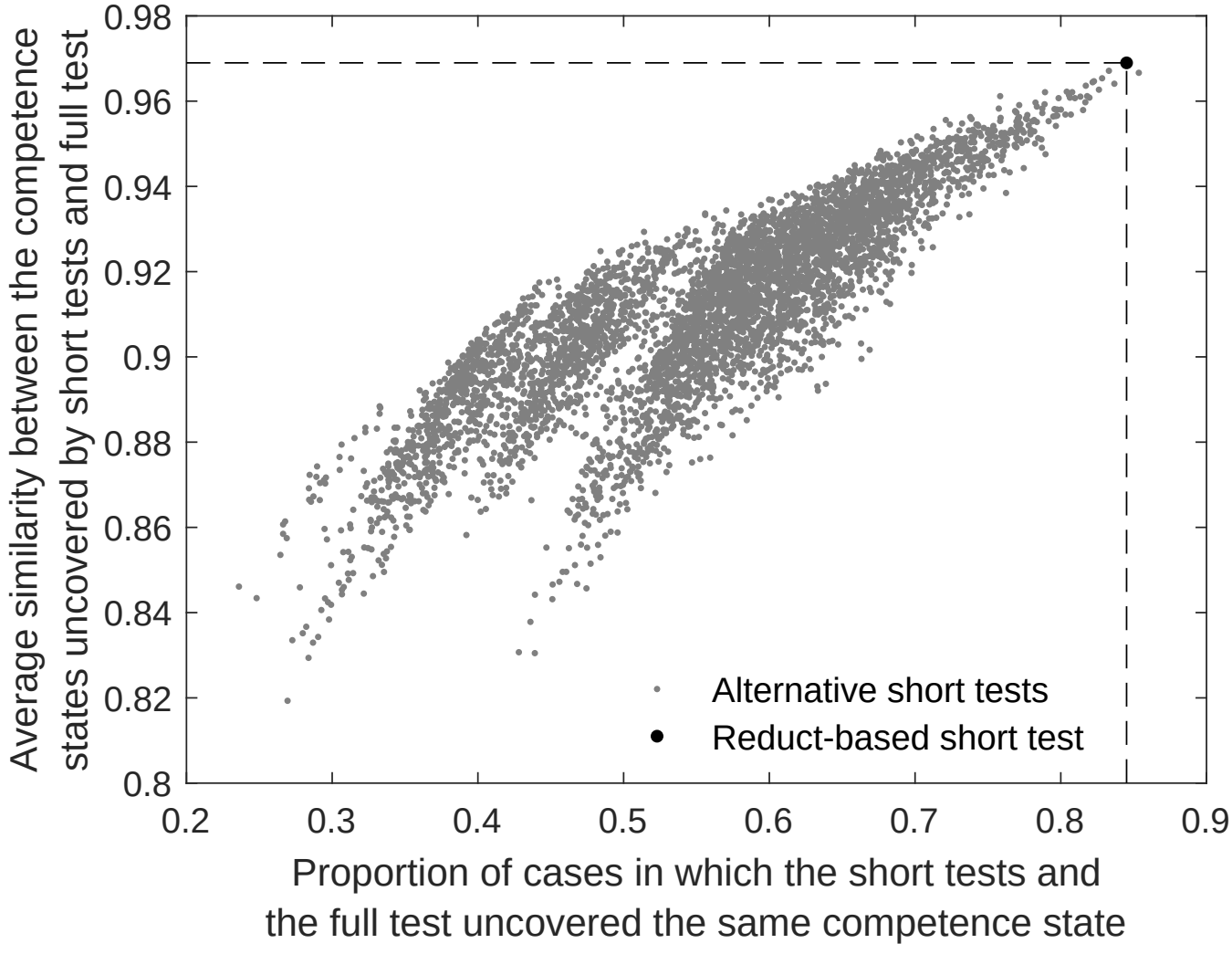
Example of shortening an existing test (III)

It is sufficient to use the following 9 items to obtain a test that is as informative as the original 20-item test

Item		Competency	Item		Competency
1	$\frac{5}{3} - \frac{3}{4}$	$\{d, f, g\}$	11	$4\frac{1}{3} - 2\frac{4}{3}$	$\{b, e, g\}$
2	$\frac{3}{4} - \frac{3}{8}$	$\{d, g\}$	12	$1\frac{1}{8} - \frac{1}{8}$	$\{g, h\}$
3	$\frac{5}{6} - \frac{1}{9}$	$\{d, g\}$	13	$3\frac{3}{8} - 2\frac{5}{6}$	$\{b, d, e, g\}$
4	$3\frac{1}{2} - 2\frac{3}{2}$	$\{b, c, e, g\}$	14	$3\frac{4}{5} - 3\frac{2}{5}$	$\{b, g\}$
5	$4\frac{3}{5} - 3\frac{4}{10}$	$\{b, d, g, h\}$	15	$2 - \frac{1}{3}$	$\{a, g\}$
6	$\frac{6}{7} - \frac{4}{7}$	$\{g\}$	16	$4\frac{5}{7} - 1\frac{4}{7}$	$\{b, g\}$
7	$3 - 2\frac{1}{5}$	$\{a, b, g\}$	17	$7\frac{3}{5} - \frac{4}{5}$	$\{b, e, g\}$
8	$\frac{2}{3} - \frac{2}{3}$	$\{g\}$	18	$4\frac{1}{10} - 2\frac{8}{10}$	$\{b, e, f, g\}$
9	$3\frac{7}{8} - 2$	$\{b\}$	19	$4 - 1\frac{4}{3}$	$\{a, b, c, e, g\}$
10	$4\frac{4}{12} - 2\frac{7}{12}$	$\{b, e, g, h\}$	20	$4\frac{1}{3} - 1\frac{5}{3}$	$\{b, c, e, g\}$

Example of shortening an existing test (IV)

The short test outperformed 5,004 alternative short tests of the same length in uncovering the competence state uncovered by the full test



Take home message

CbTD is a useful and usable method for constructing, improving, and shortening tests that assess the skills an individual masters based on item responses

It identifies the exact skill combinations items must require for the test to be as informative as possible and minimal

While the development of conjunctive educational tests has been illustrated here, CbTD also supports:

- disjunctive tests – any skill associated with an item is sufficient to solve it
- clinical tests – symptoms of a disorder are assessed based on item responses