Chapter 6:

Mastery learning

EPFL

How do people learn ?

☑ by exploration, trial and error

✓ **by incremental mastery**

- by verbal elaboration

CS411 Fall 2025 Assignment and checkpoint due dates

	Exercises **Tuesday, 08:15 - 10:00** **MED 0 1418**	**Lectures** **Tuesday, 10:15 - 12:00** **MED 0 1418**		**Due dates for check points and assignments**
09.09	–	Overview of learning technologies	P. Dillenbourg & T. Käser	
16.09	Experimenting PS-I Defining learning goals	Human cognition	Pierre Dillenbourg	**Checkpoint 1 Due 18.09 (Thursday 23h59)** Form groups, choose topics, and submit on Moodle
23.09	Designing learning activities	Designing experiments	Chris Petrie	**Checkpoint 2 Due 25.09 (Thursday 23h59):** Submit 1-2 page description of problem-solving activity and learning goals
30.09	Designing learning activities	Piaget, microworlds & simulations	Pierre Dillenbourg	
07.10	Designing learning activities	Inferential statistics (part 1)	Patrick Jermann	
14.10	Implementing learning activities	Inferential statistics (part 2)	Patrick Jermann	**17.10 Milestone 1 (Pass)** Submit 3-4 page summary of project plan including description of problem-solving activity and sketch.
~~21.10~~	*Break*			
28.10	Implementing learning activities	Skinner, Bloom & mastery learning	Pierre Dillenbourg	
04.11	Statistics	Student modeling, Adaptive Ed.	Tanja Käser	
11.11	Statistics	Bayesian Knowledge Tracing	Tanja Käser	
18.11	Running Experiments	Vygotsky, social cognition	Pierre Dillenbourg	
25.11	Running Experiments	GenAI for Education	Tanja Käser	
02.12	Running Experiments	GenAI for Education	Tanja Käser	
09.12	Analyzing data	AR & VR For education	Pierre Dillenbourg	
16.12	Finalizing projects	Synthesis & project presentation	P. Dillenbourg & T. Käser	**Submit final report 06.01.2026 (Tuesday)**

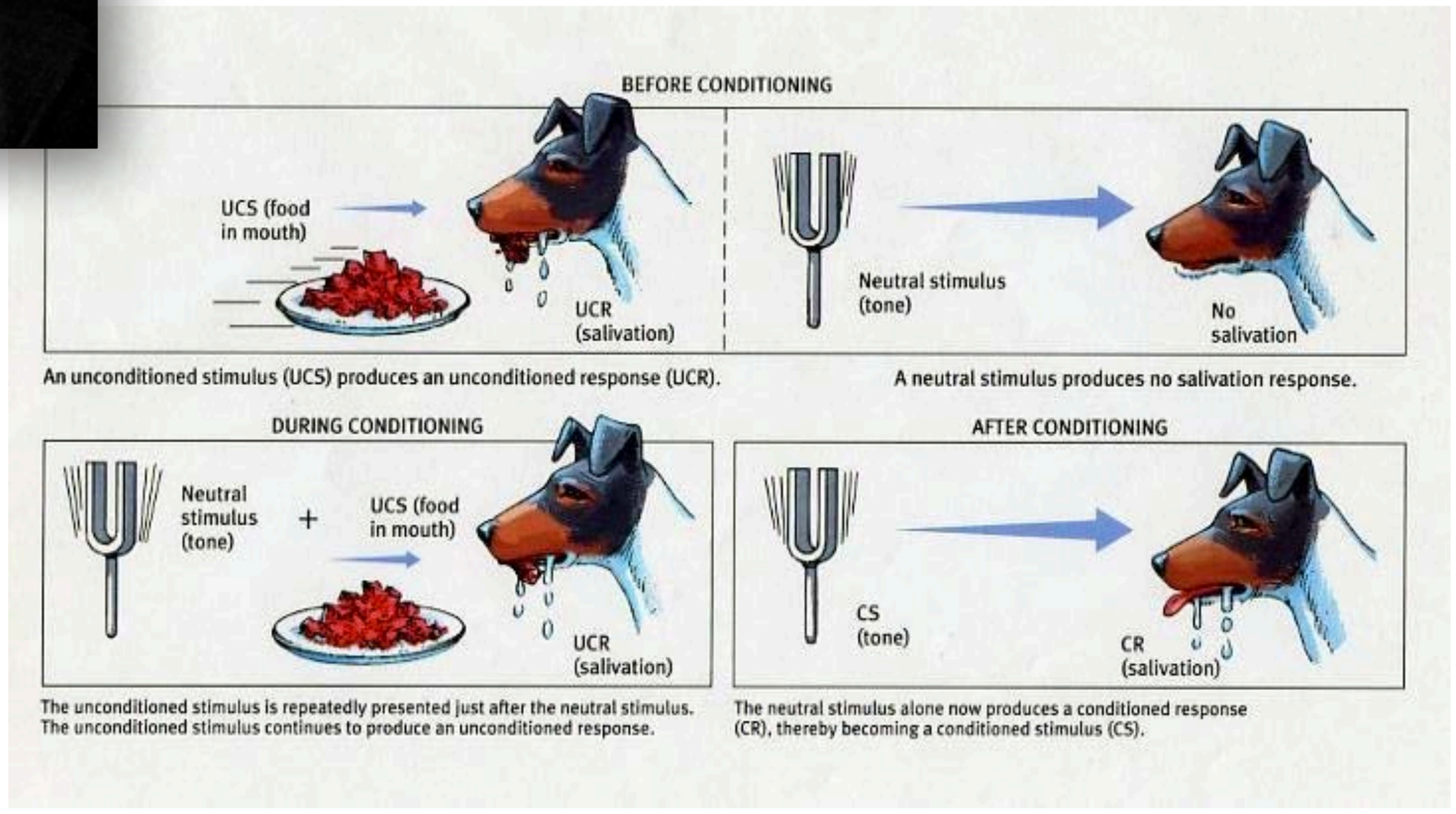
dogs

How do people learn ?

http://www.nobelprize.org/nobel_prizes/medicine/laureates/1904/pavlov-bio.html

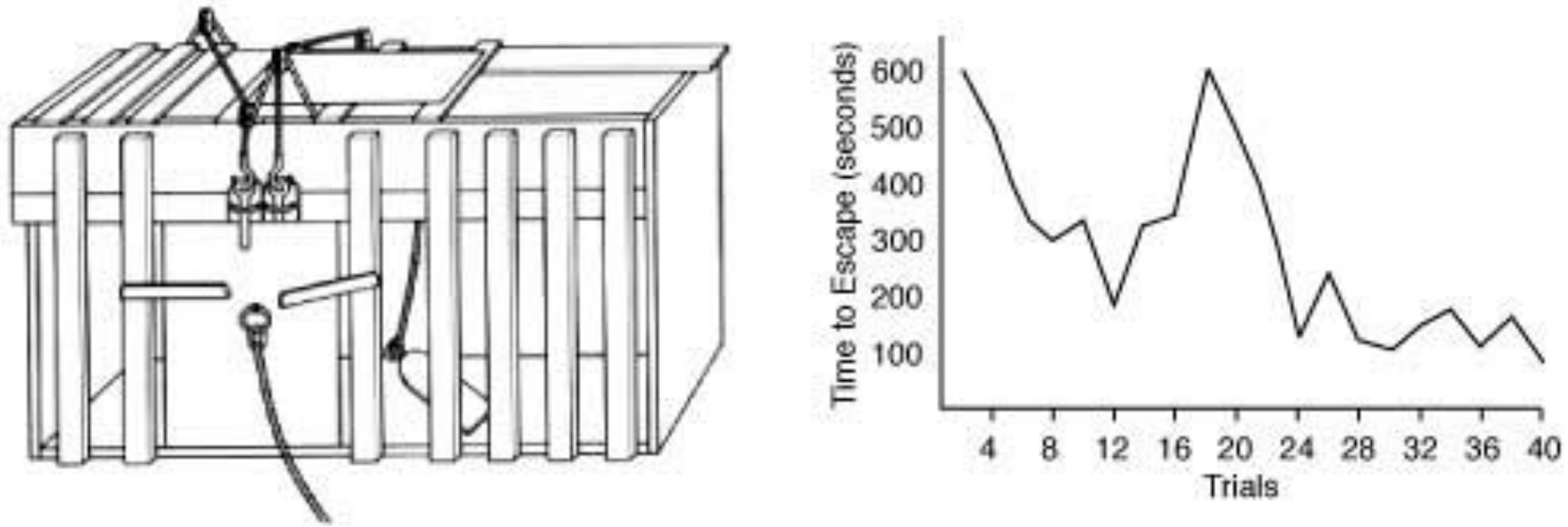
Yvan Pavlov, 1849-1936

Classical Conditioning



http://www.rhsmpsychology.com/Handouts/classical_conditioning.htm

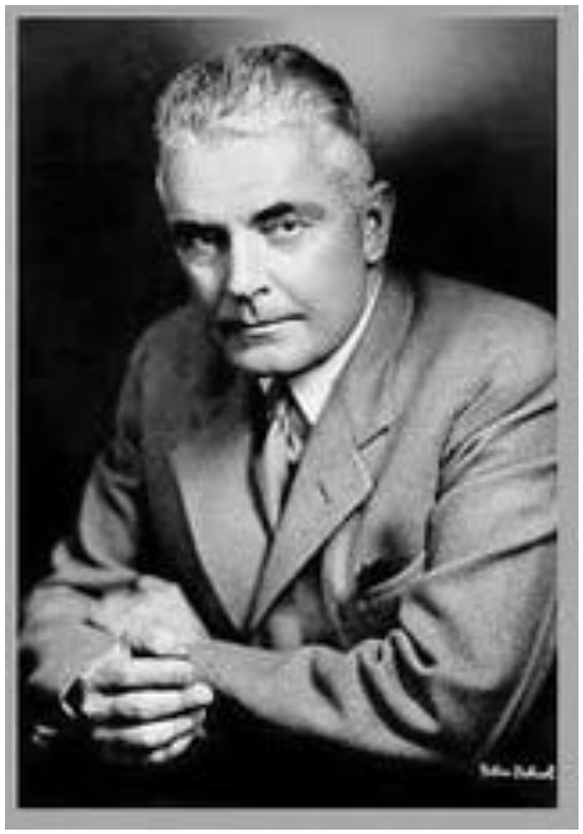
Edward L. Thorndike (1874 – 1949) The Law of Effect : any behavior that is followed by pleasant consequences is likely to be repeated, and any behavior followed by unpleasant consequences is likely to be stopped.



Adapted from Domjan, 1993 (modified from Thorndike, 1898 [left] and Imada & Imada, 1983 [right])

http://www.simplypsychology.org/edward-thorndike.html

Edward L. Thorndike, The Law of Effect, The American Journal of Psychology
Vol. 39, No. 1/4 (Dec., 1927), pp. 212-222: http://www.jstor.org/stable/1415413



http://psychology.about.com/od/behavioralpsychology/f/behaviorism.htm

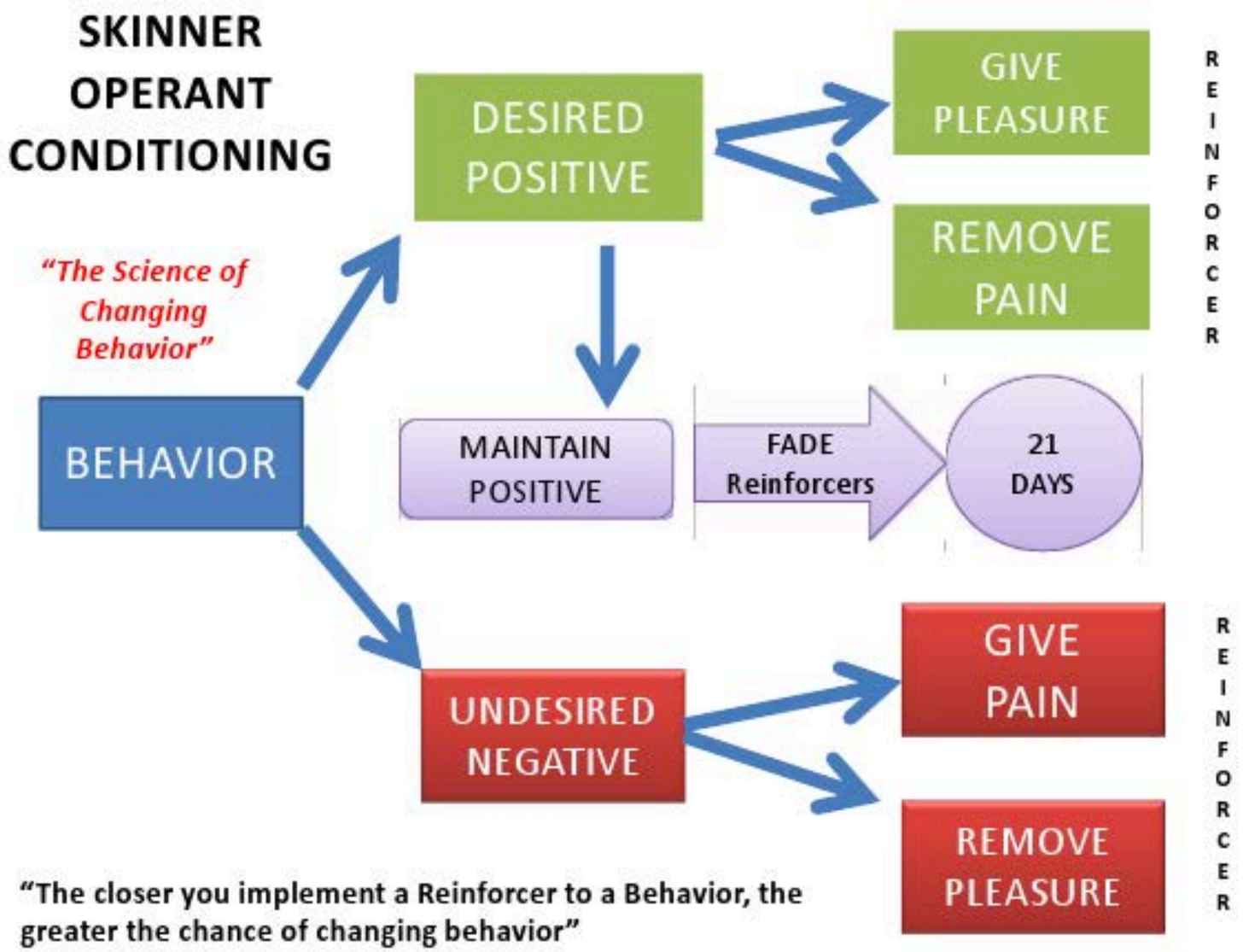
John Watson (1878- 1958) Behaviourism

"Give me a dozen healthy infants, well-formed, and my own specified world to bring them up in and I'll guarantee to take any one at random and train him to become any type of specialist I might select -- doctor, lawyer, artist, merchant-chief and, yes, even beggar-man and thief, regardless of his talents, penchants, tendencies, abilities, vocations, and race of his ancestors."

--John Watson, Behaviorism, 1930

The Little Albert Experiment
https://www.youtube.com/watch?v=5duLMjaTL0U

Burrhus Frederic Skinner (1904-1990), Operant Conditioning
https://www.youtube.com/watch?v=NeK8GNLylkc



https://sites.google.com/a/adams12.org/harp-mrhs-home-v1/06-powerpoint-section/skinner-s-operant-conditioning

Key ideas in behaviorism

① Psychology is becoming more scientific

② The brain is a black box; the focus is on behaviors

③ Learning is « engineered »

④ Association results from immediate feedback

⑤ The learner is permanently active

⑥ Small steps increase the probability of positive feedback → Programmed instruction

B. F. Skinner

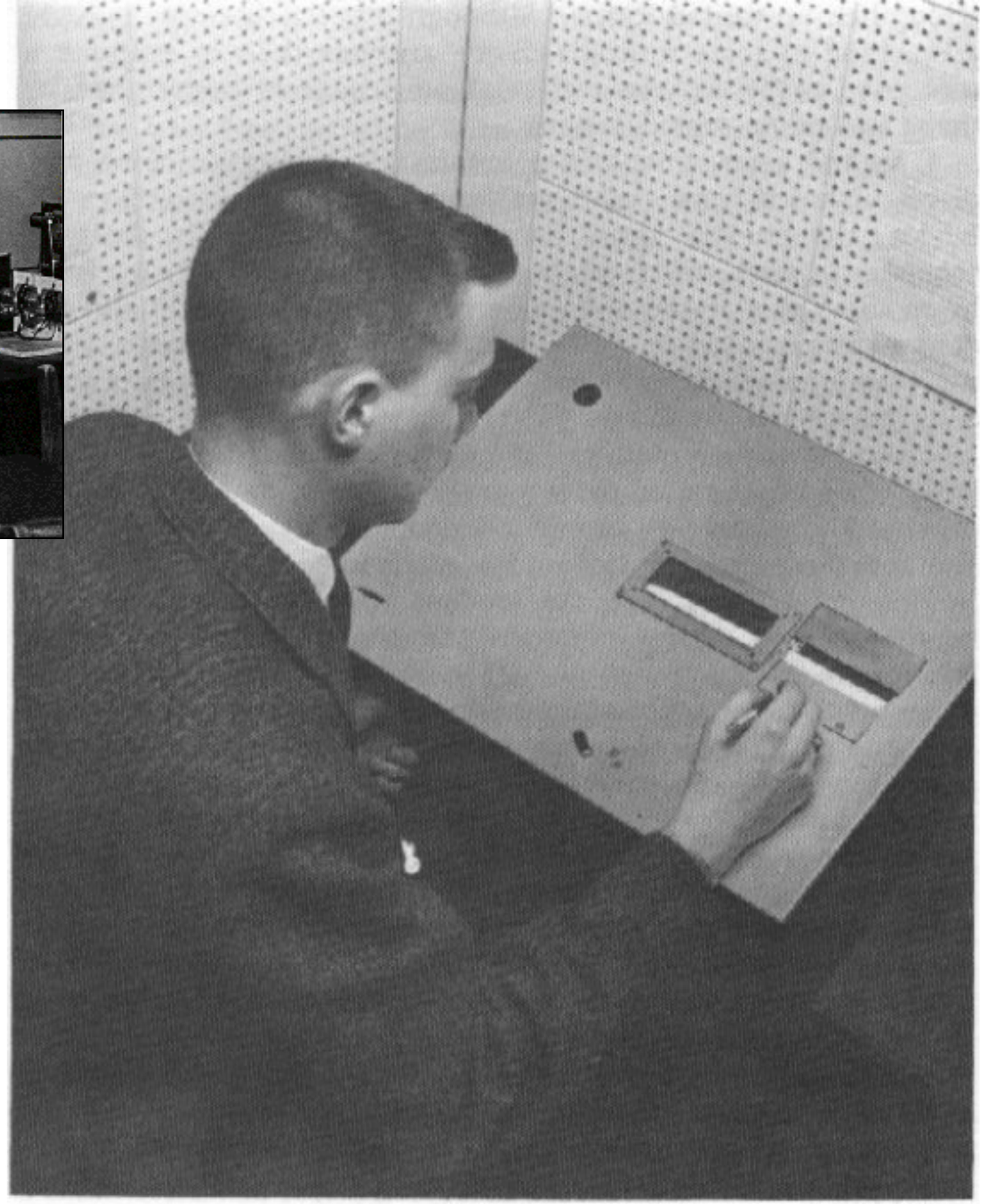
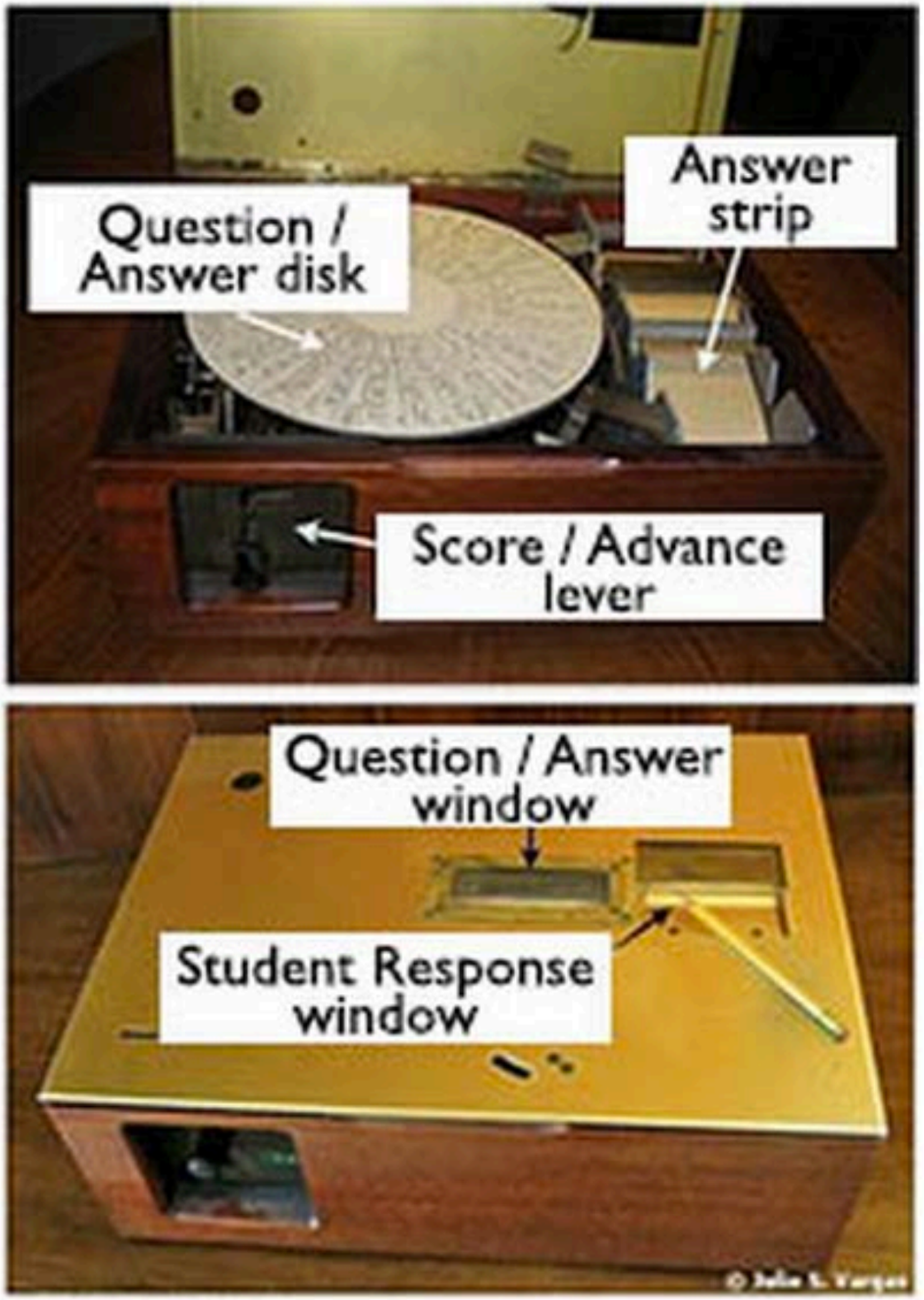


FIG. II. Student at work in the self-instruction room. Material appears in the left-hand window. The student writes his response on a strip of paper exposed at the right.

B.F. Skinner's Teaching Machine



(c) Julie S.Vargas

Instructional "disks" are placed inside the machine along with a strip or roll of paper. When the machine is closed, the student reads a question through a window and writes their response on the strip of paper. The student then compares their answer with the answer on the disk and presses the lever one way if their answer is correct or the other way if incorrect (the machine keeps score and advances).

Skinner proposed the machine improves learning by "taking into account the rate of learning for each individual learner." With this, Skinner formalizes "self-paced instruction" as part of programmed instruction.

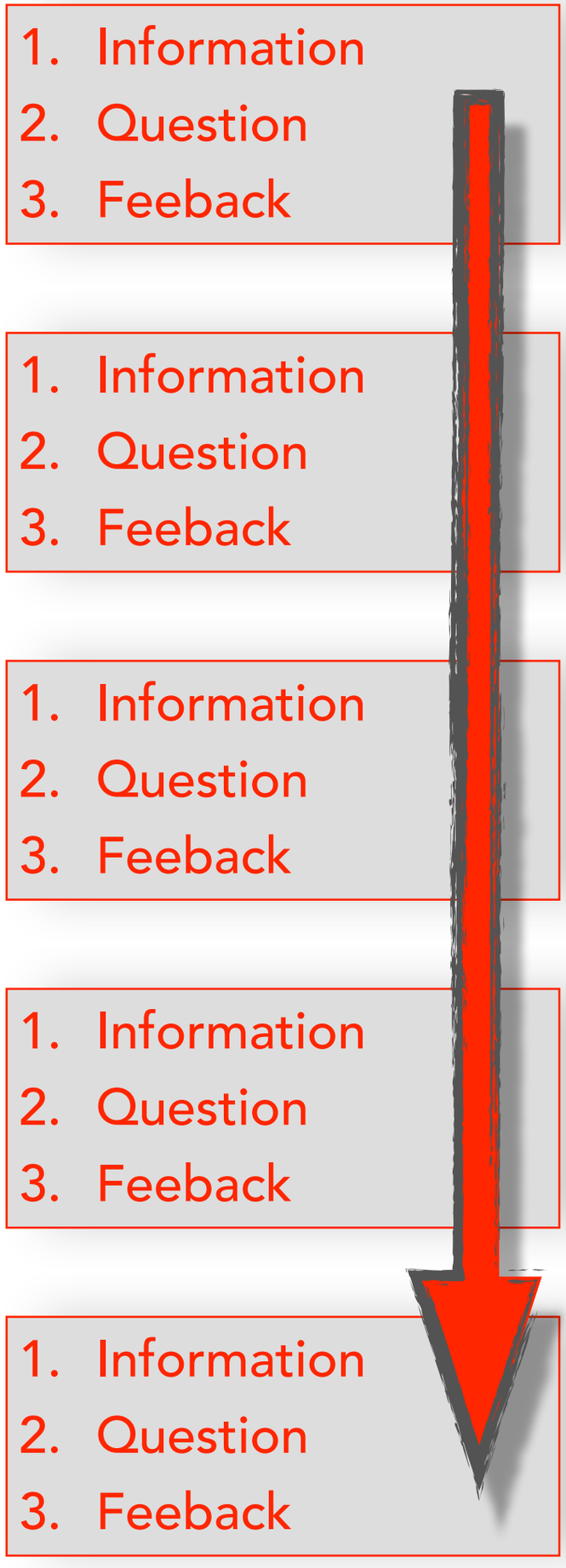
Linear Instruction

Table 2. PART OF A PROGRAM IN HIGH-SCHOOL PHYSICS

The machine presents one item at a time. The student completes the item and then uncovers the corresponding word or phrase shown at the right.

SENTENCE TO BE COMPLETED	WORD TO BE SUPPLIED
1. The important parts of a flashlight are the battery and the bulb. When we "turn on" a flashlight, we close a switch which connects the battery with the _____.	bulb
2. When we turn on a flashlight, an electric current flows through the fine wire in the _____ and causes it to grow hot.	bulb
3. When the hot wire glows brightly, we say that it gives off or sends out heat and _____.	light
4. The fine wire in the bulb is called a filament. The bulb "lights up" when the filament is heated by the passage of a(n) _____ current.	electric
5. When a weak battery produces little current, the fine wire, or _____, does not get very hot.	filament
6. A filament which is **less** hot sends out or gives off _____ light.	less
7. "Emit" means "send out." The amount of light sent out, or "emitted," by a filament depends on how _____ the filament is.	hot

http://66.31.107.147/~heines/academic/papers/2002elearn/images/Lumsdaine_Glaser/



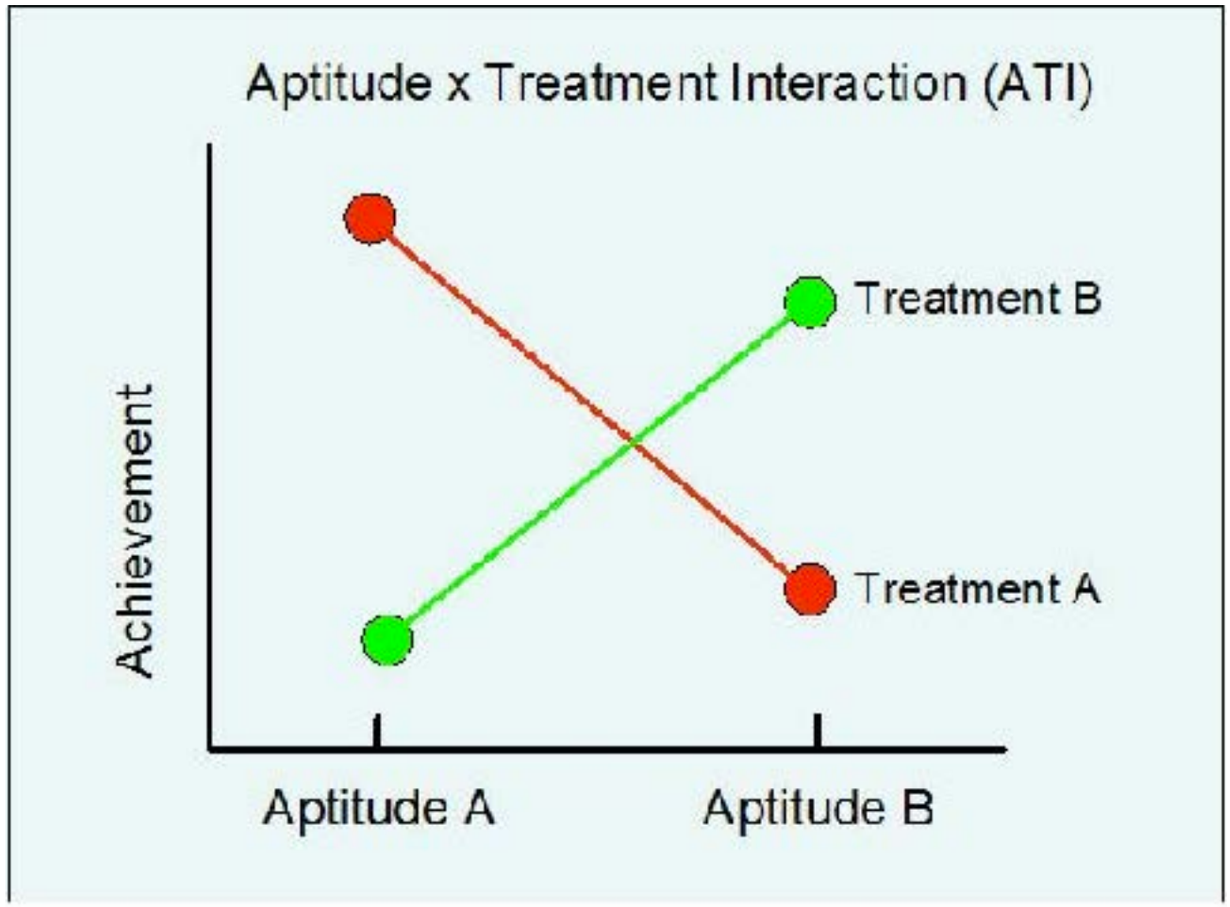
Frame-Based Models / e-learning

1. Decomposition: Segmenting complex contents into a sequence of learning steps that contains an elementary piece of information
2. Keep the student active all the time, ask the student to process any new piece of information
3. Provide immediate feedback
4. Let the student move on at his or her own speed

Feedback

is the 1st principle

for pedagogical effectiveness



Aptitude- Treatment Interactions:

The effect of a pedagogical method varies for different learners profiles

Cronbach, L. & Snow, R. (1977). Aptitudes and Instructional Methods: A Handbook for Research on Interactions. New York: Irvington

The effect of a pedagogical method

varies for different learners profiles

- Personalized Instruction
- Individual instruction
- Adaptive instruction
- ...
- *Optimization in learning environments*

Evaluations of Intelligent Tutoring Systems

- Study with 17,000 students showed that Cognitive Tutor Algebra (a curriculum + ITS) doubled students' algebra learning

  (Pane et al., 2013)

- Meta-review indicates that ITSs are "nearly as effective as human tutoring" (VanLehn, 2011)

- Four meta-analyses show ITSs are often more effective than other forms of instruction

  (Kulik & Fletcher, 2015; Ma, Adesope, Nesbit, & Liu, 2014; Steenbergen-Hu & Cooper, 2013; 2014)

By courtesy of Vincent Aleven, CMU

Instructional design starts with :

~~What should learners know at the end ?~~
~~(which they did not know at the beginning)~~

What should learners be able to do at the end ?
(which they could not do at the beginning)

How will I know they are able to do it ?

What would be the exam questions or tasks ?

What should learners <u>be able to do</u> at the end ?

Pedagogical Objectives

Learning Goals

Learning Outcomes

Reproduction

Conceptualisation

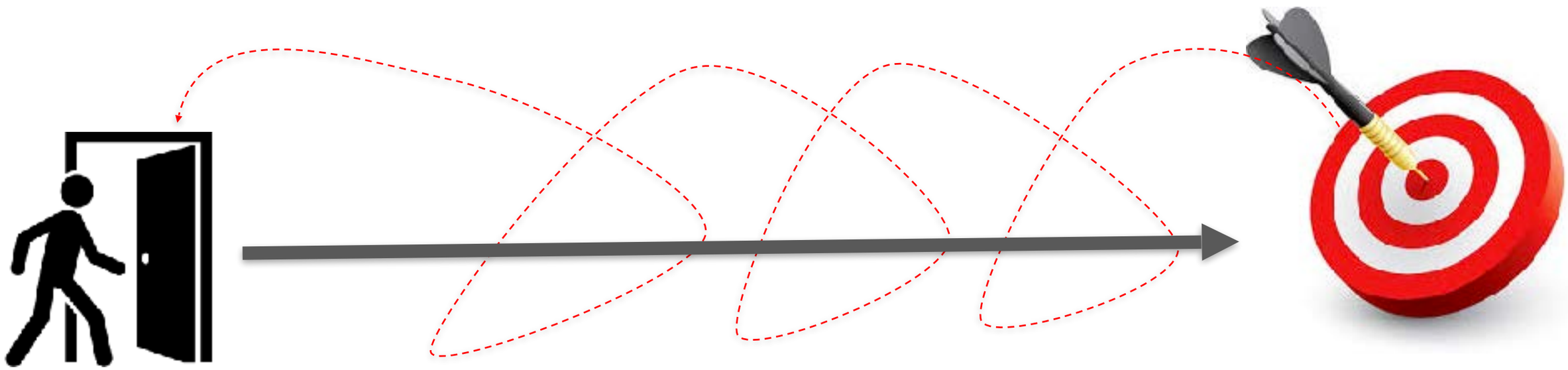
Application

Exploration

Mobilisation

Problem solving

Instructional design works backwards :



What are they able to do at the beginning

Pre-Requisites

What should learners be able to do at the end ?

(which they could not do at the beginning)

Objectives

Skill1 is pre-requisite to Skill2
if p(Skill2 |~Skill) << p(Skill2 | Skill1)

Skill1 : « 6 + 9 »
Skill2: « 26 + 39 »

Skill1: to find the verb
Skill2: to agree it with the subject

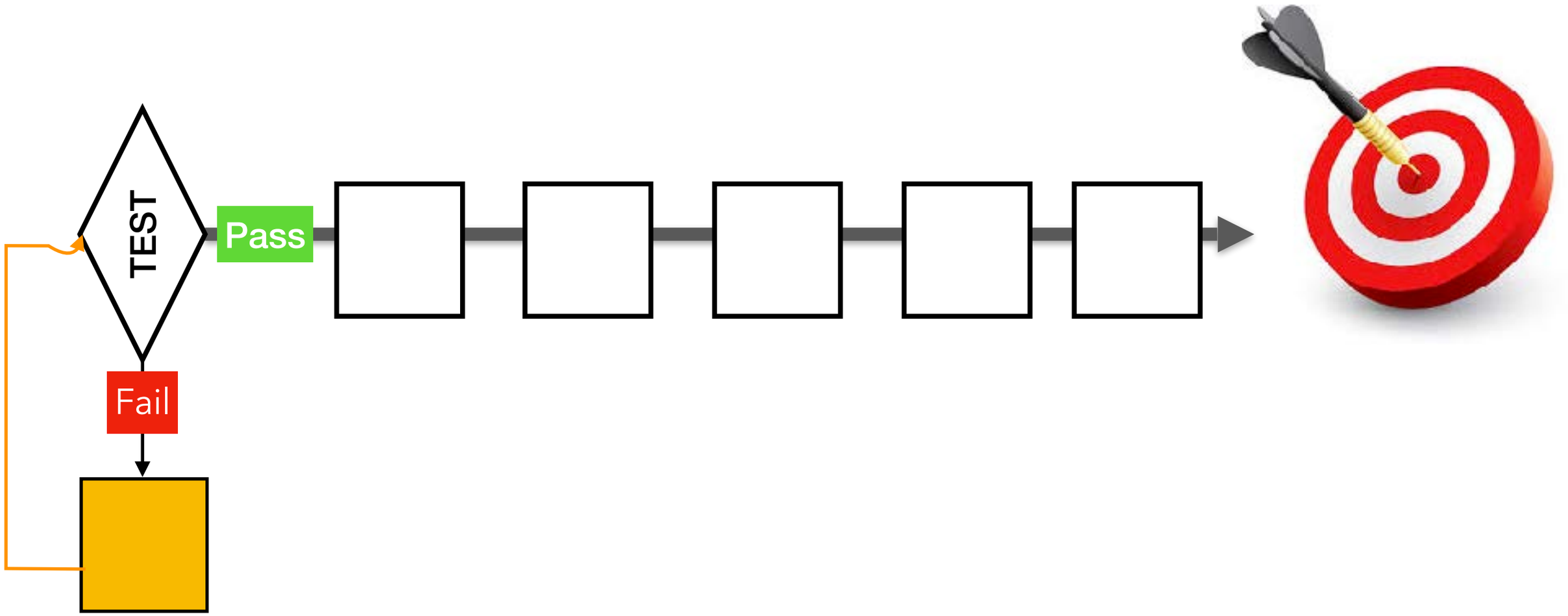
Skill1 : angle
Skill2: square

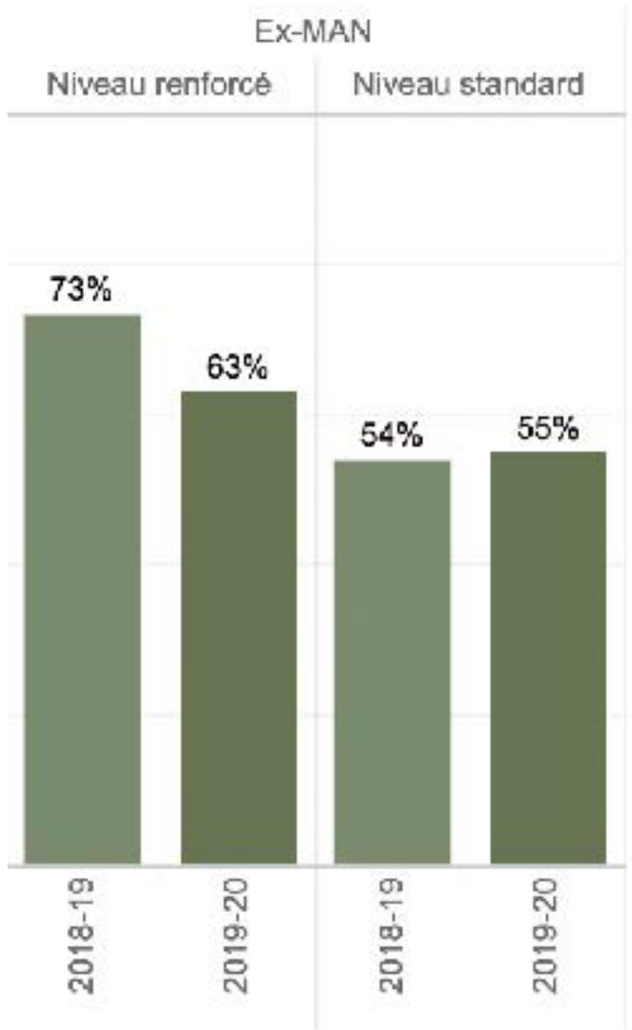
Skill1 : mean
Skill2: standard deviation

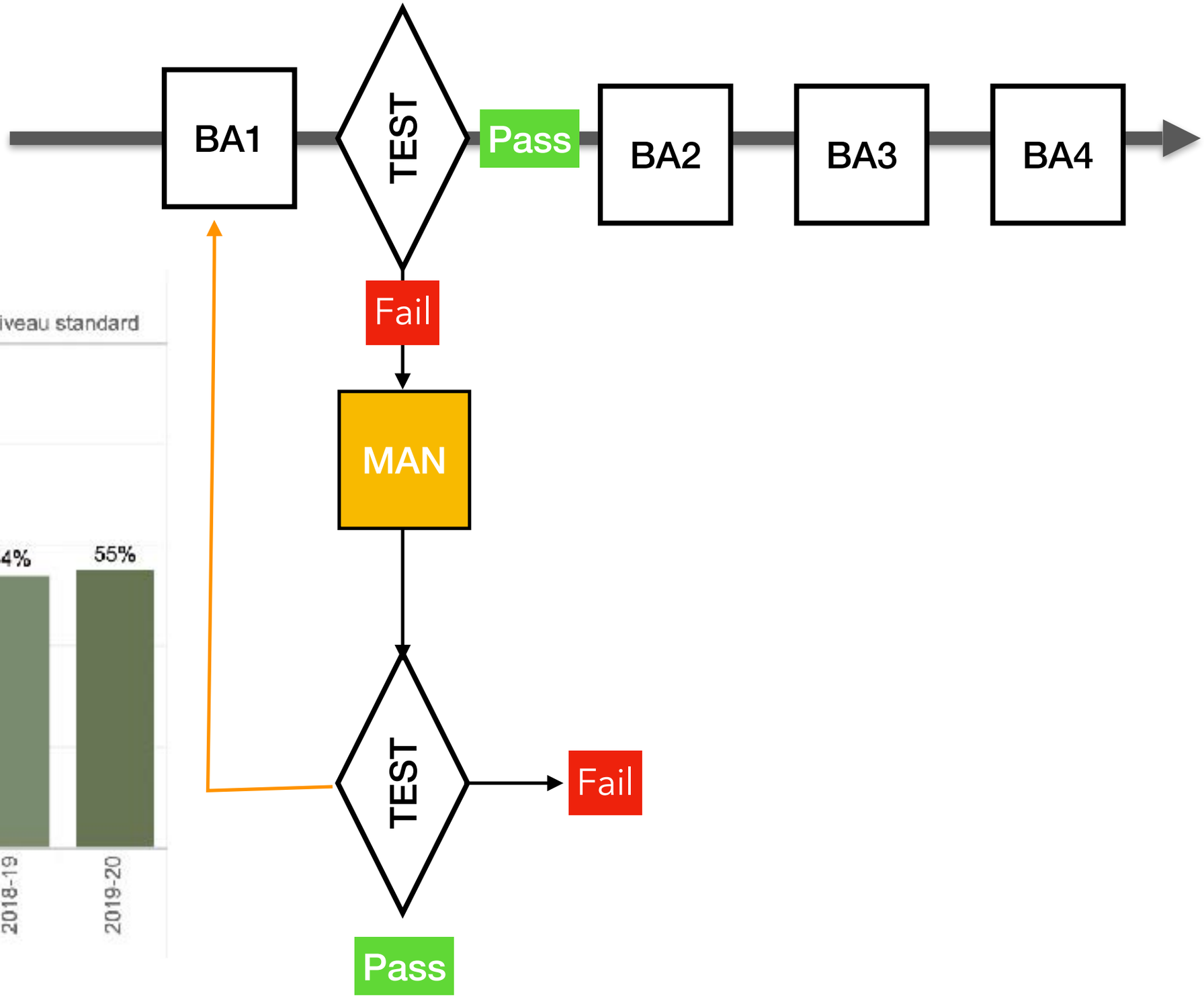
Skill1 : to read
Skill2: to code

Mastery learning:

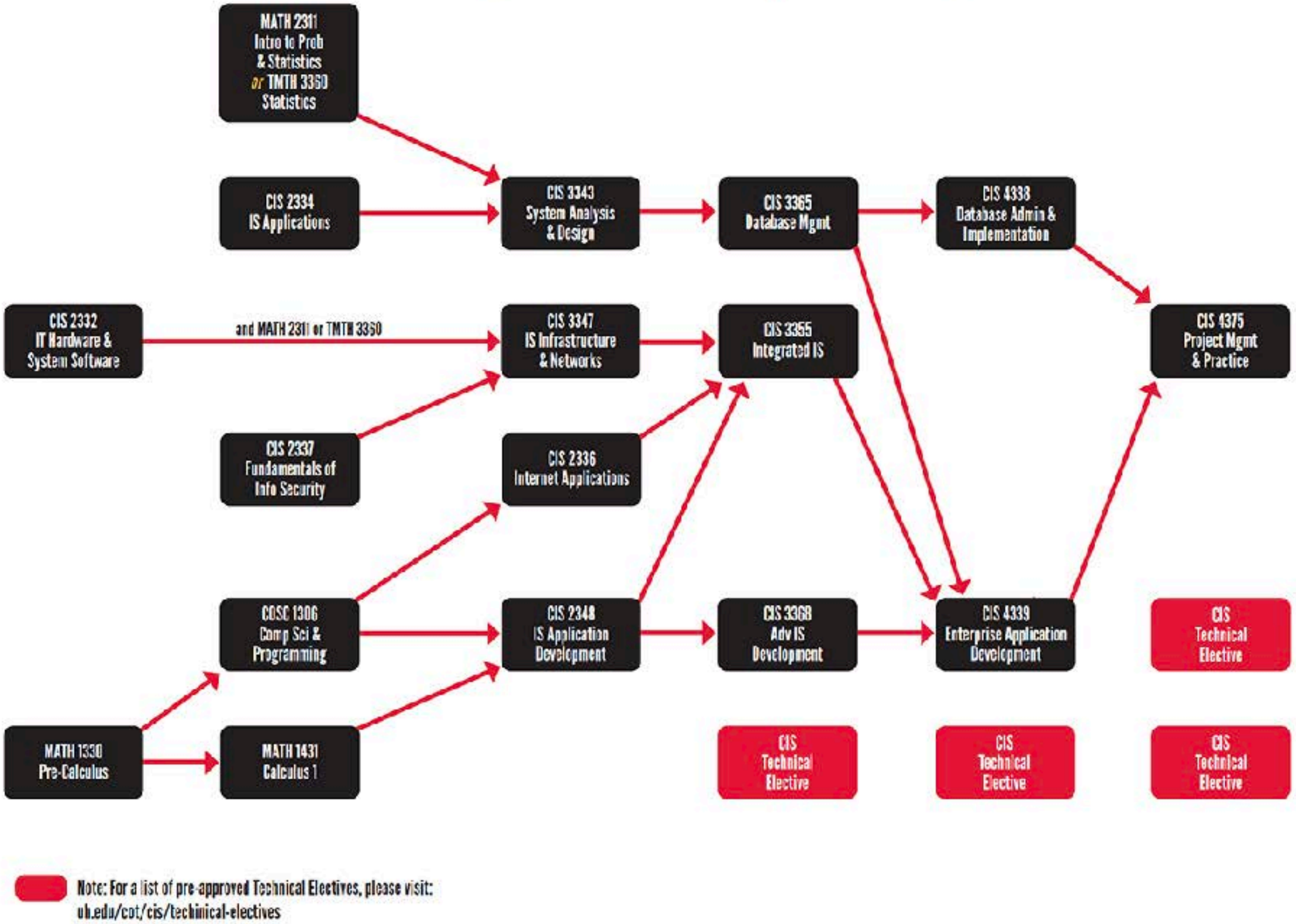
if pre-requisites are missing, remediate them before to go on




Ex-MAN
Niveau renforcé
Niveau standard
73%
63%
54%
55%
2018-19
2019-20
2018-19
2019-20

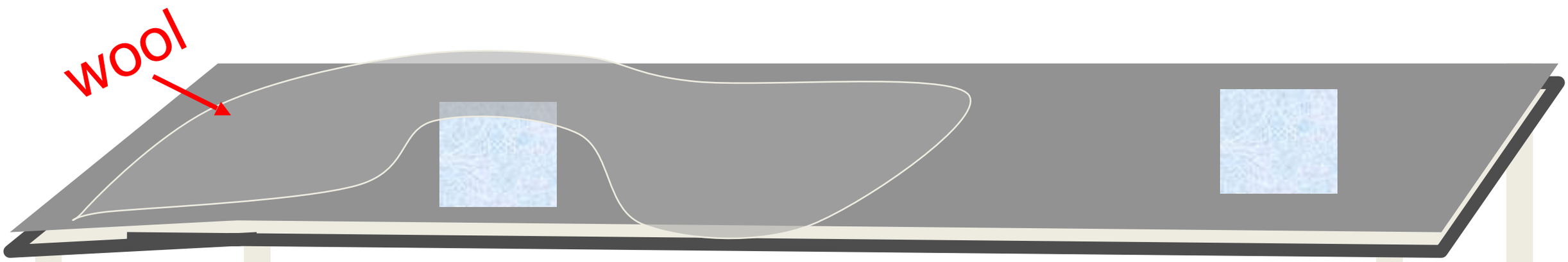

BA1
TEST
Pass
BA2
BA3
BA4
Fail
MAN
TEST
Fail
Pass

CIS Prerequisite/Course Sequence Map



https://www.uh.edu/technology/programs/undergraduate/computer-information-systems/prerequisite-map/

Prior Knowledge can be wrong



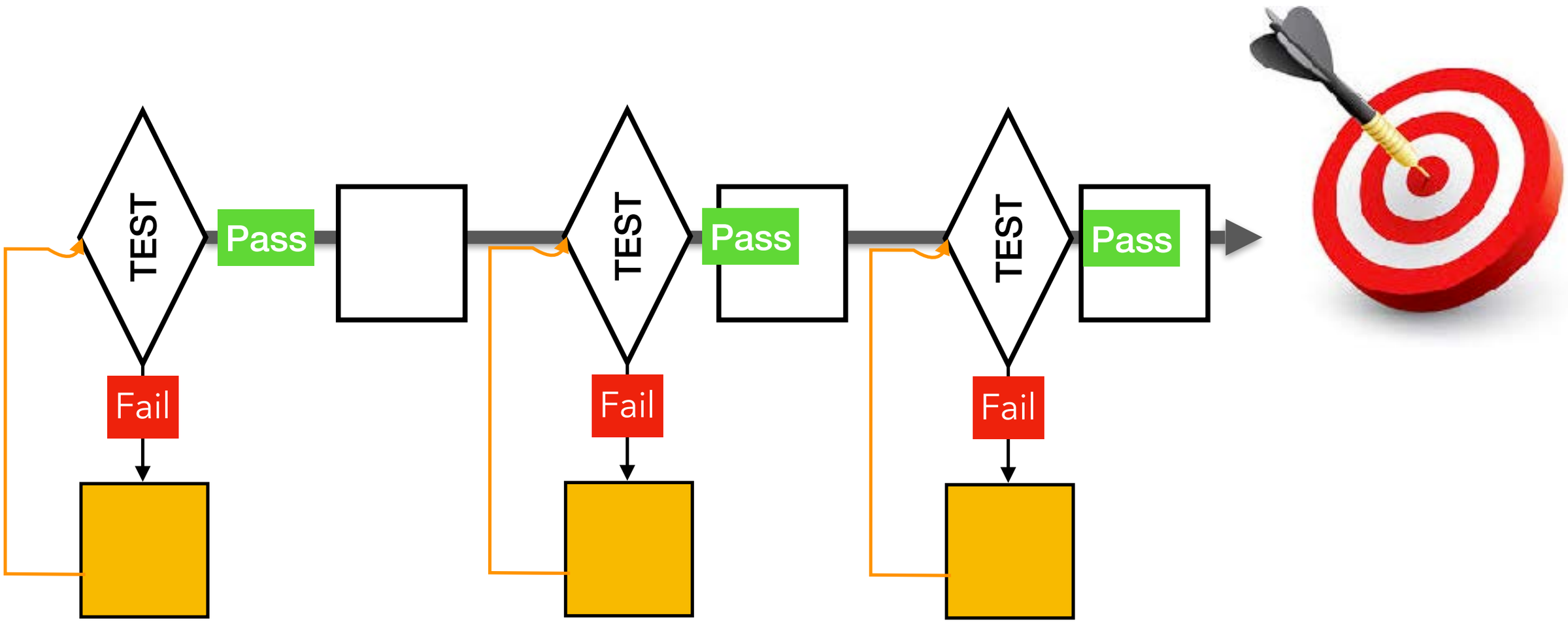
2 ice cubes at midnight, t = 20 C

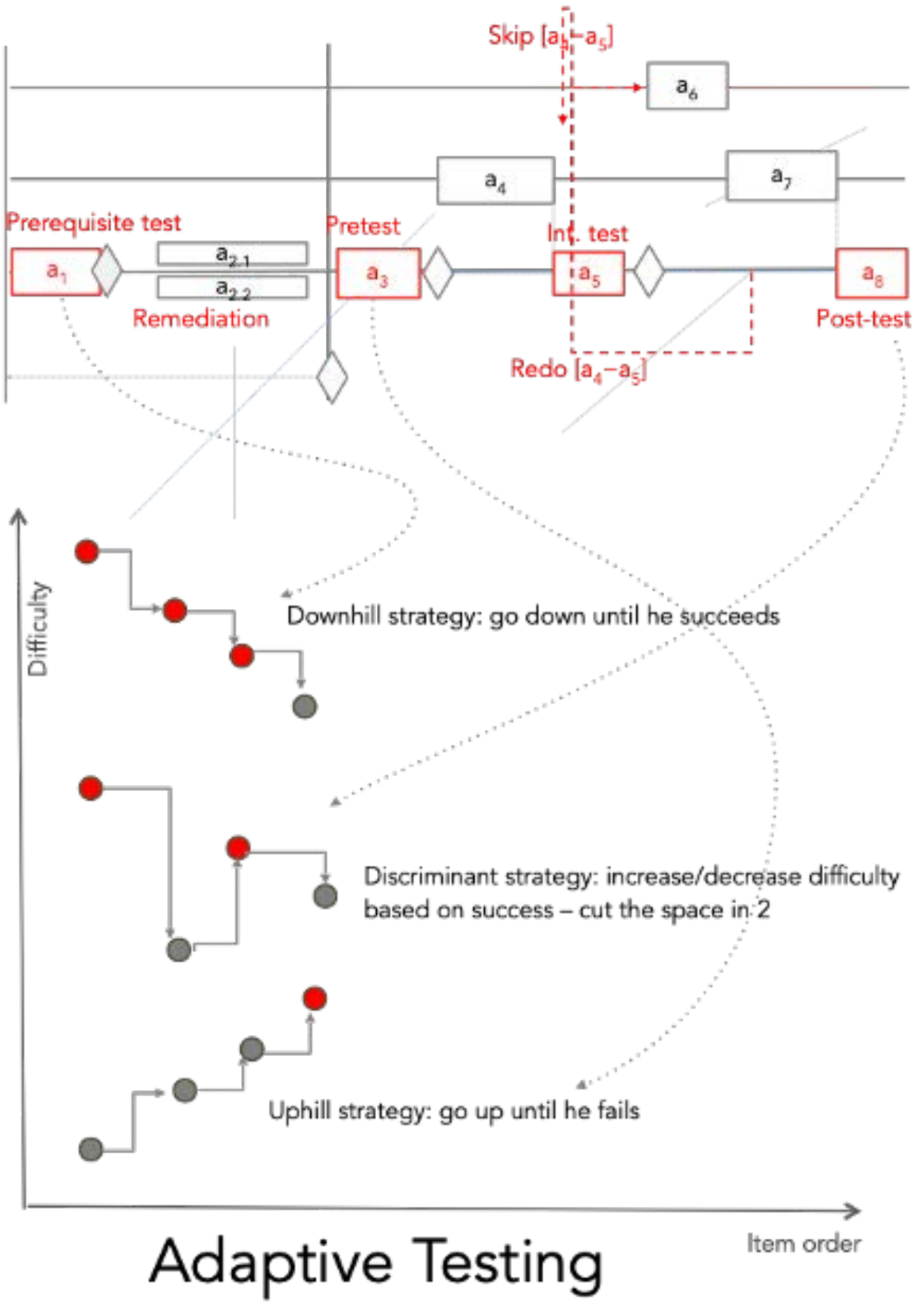
Which ice cube will melt faster ?

Misconceptions must me trapped otherwise they survive teaching

Mastery learning:

Control mastery of $skill_i$ before to go to $skill_{i+1}$





Adaptive Testing

Modular Instruction

Pre-requisite test: Does the learner has the pre-requisite to start the course ?

Pre-test: Should the learner skip some modules ?

Intermediate-test: Did the learner reach the objectives of this module ?

Post-test: Did the learner reach the objectives of this course?

(1) choose the topic

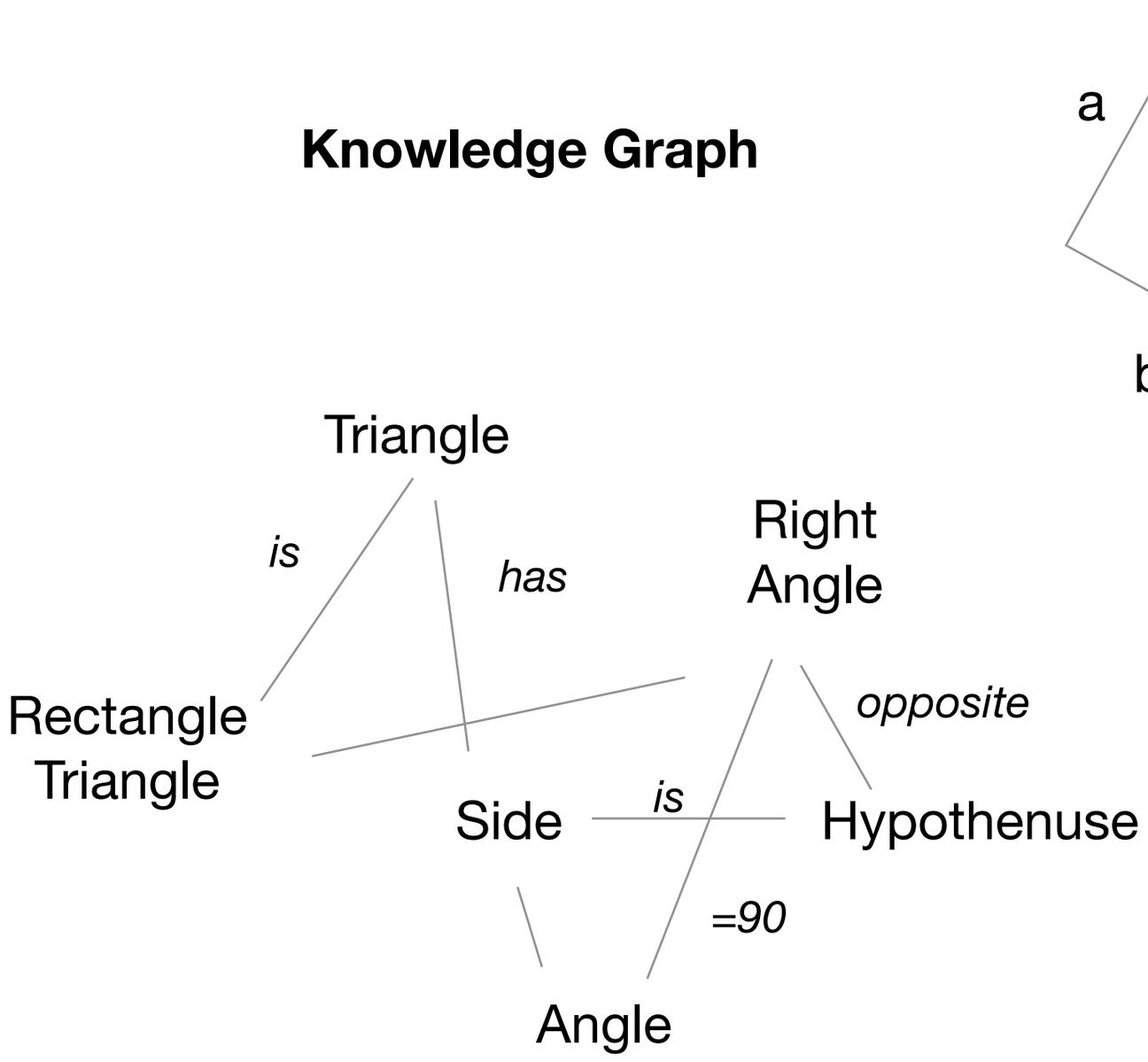
(4) Decompose

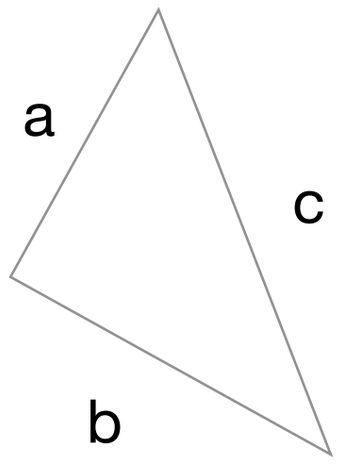
(3) Define target audience

(2) Specify Objectives

(4) Decompose

Knowledge Graph

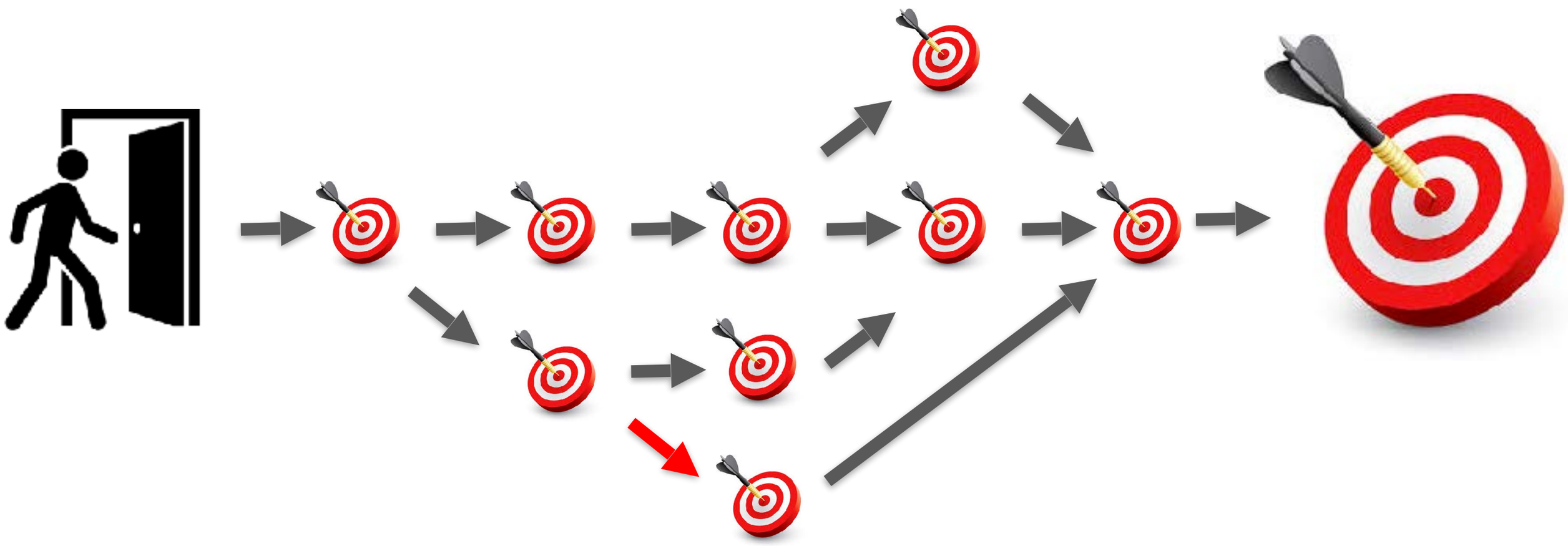




Algorithm

1. Measure A
2. Compute A^2
3. Measure B
4. Compute B^2
5. $A^2 + B^2 = C^2$
6. $C = SQRT(C^2)$

Adaptive Instruction



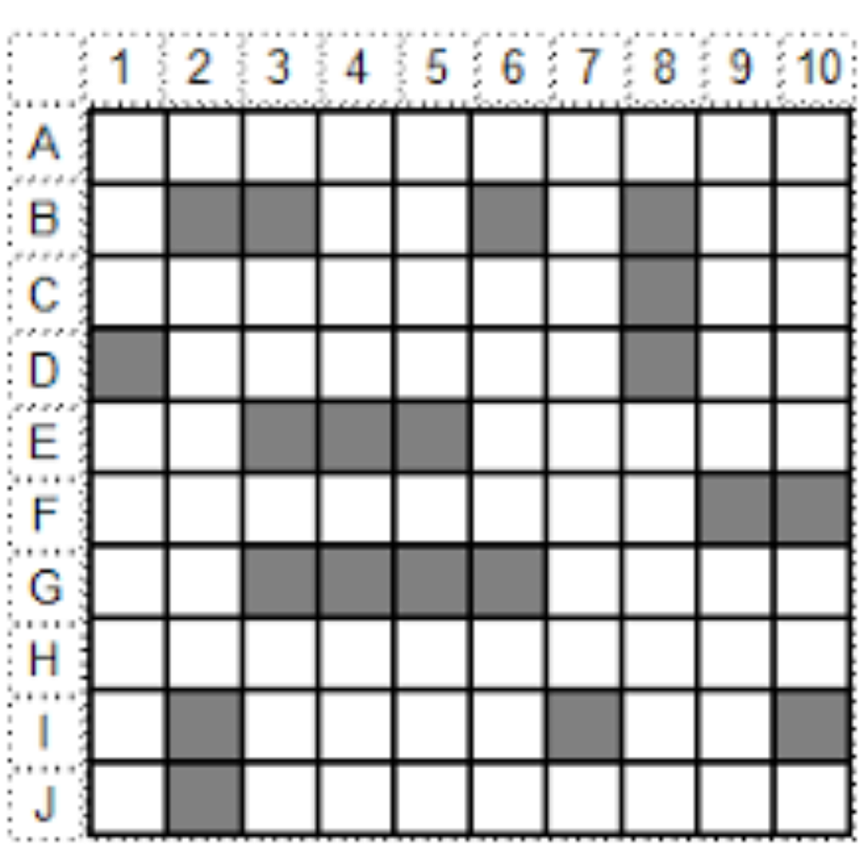
Why is an activity a_i useful for an activity a_{i+1}?

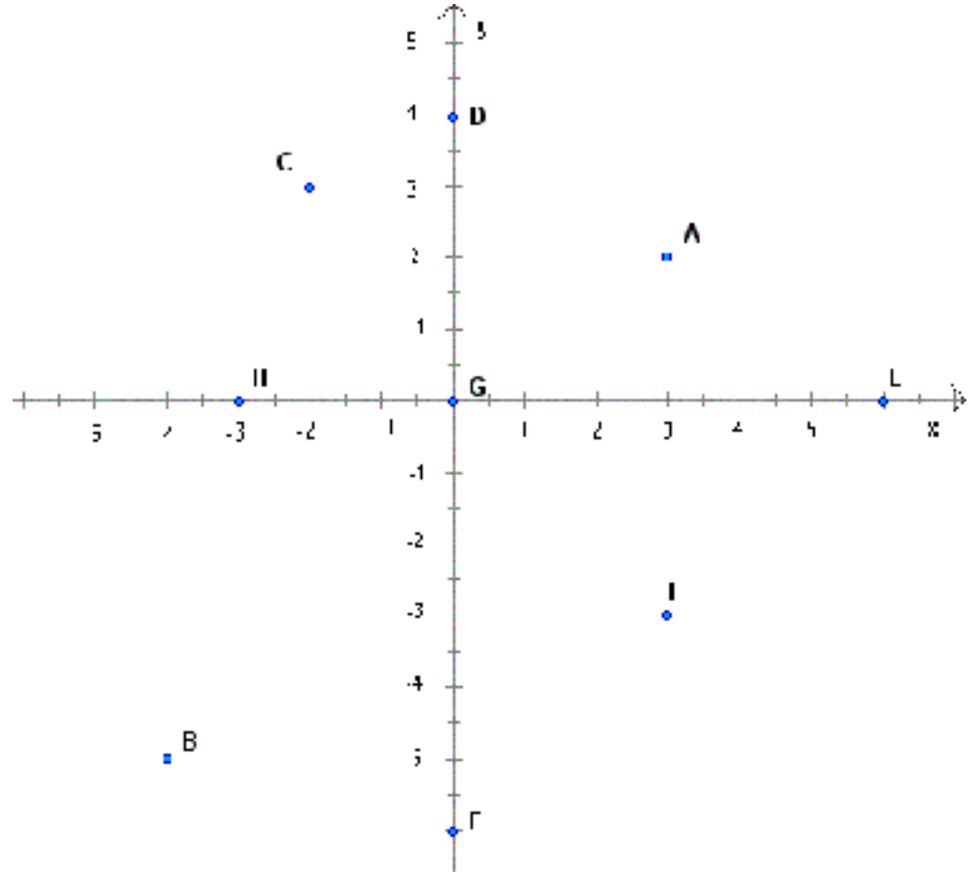
1. a_i is prerequisite to a_{i+1}
2. ...

Why is an activity a_1 useful for an activity a_2?

Advance organizer :

a_1 pre-activates structures for a_2





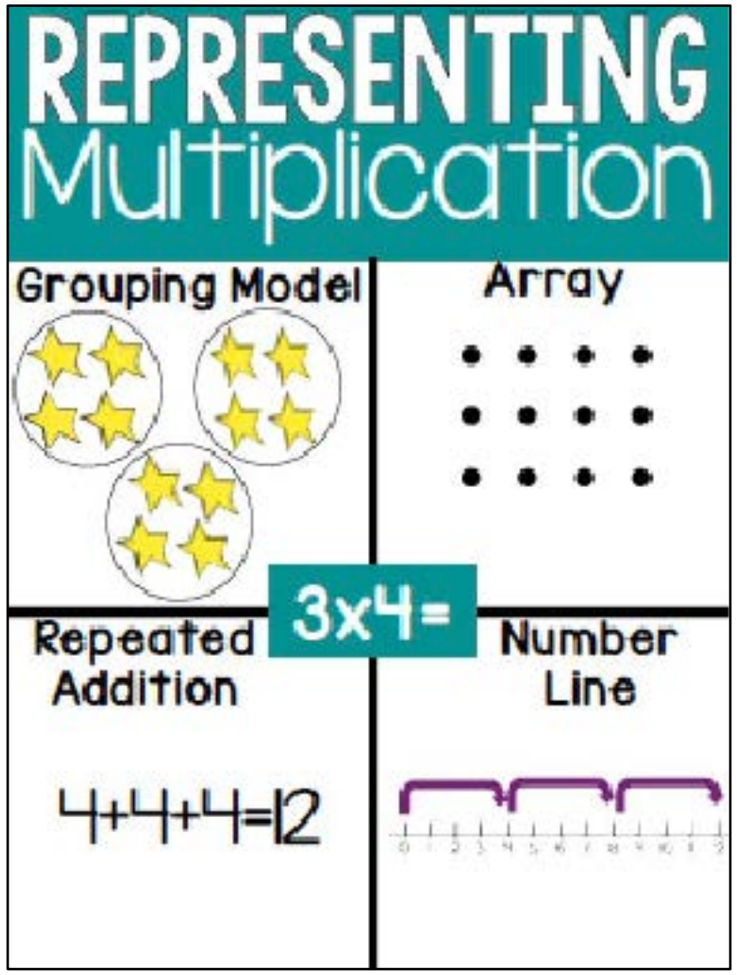
Why is an activity a_i useful for an activity a_{i+1}?

1. a_i is prerequisite to a_{i+1}
2. a_i is pre-activates cognitive schema for a_{i+1}
3. ...

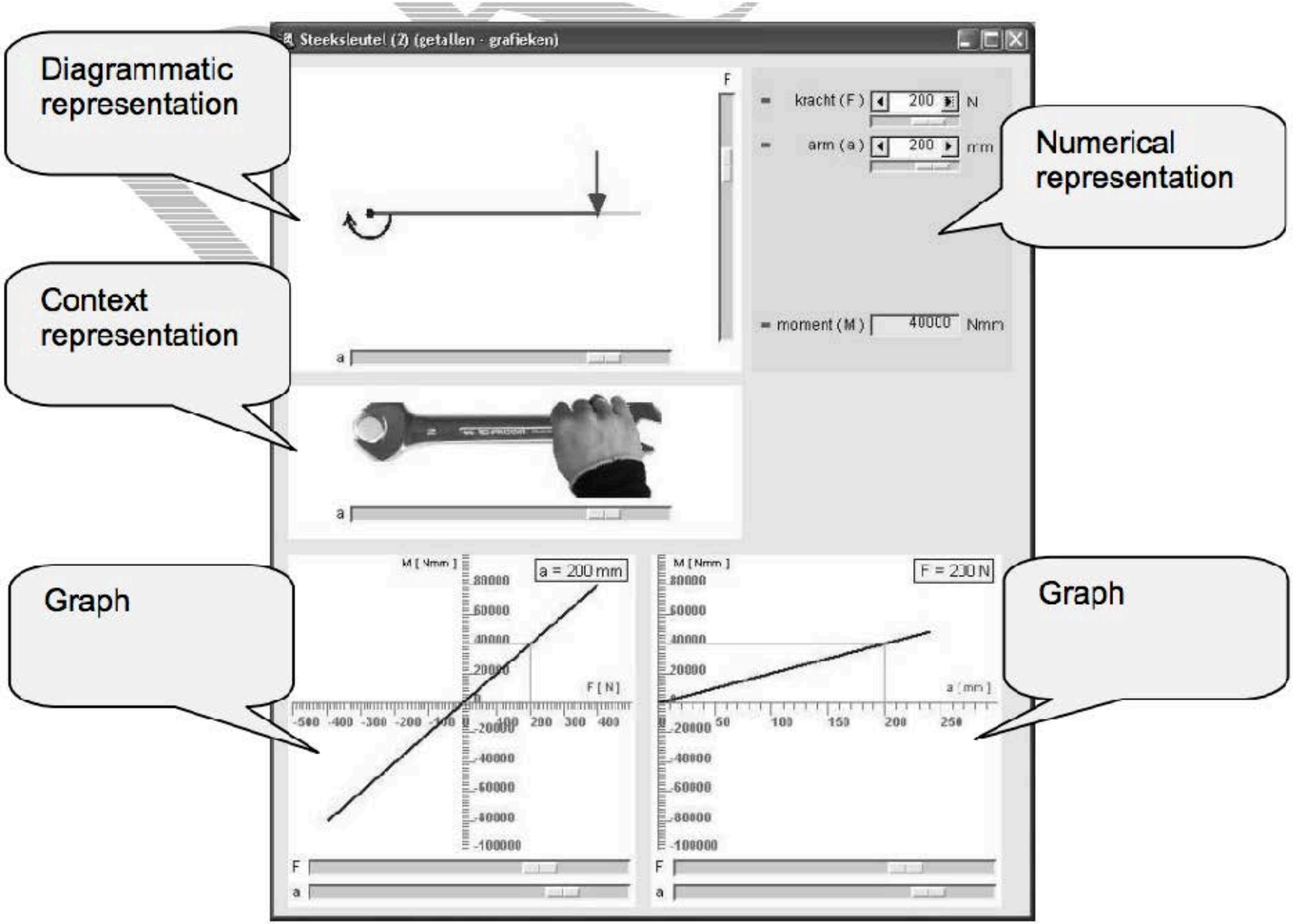
Why is an activity a_1 useful for an activity a_2?

Shift Representation:

a_2 represents contents differently from a_1



There is a large body of empirical studies that show that is beneficial for learners to switch between multiple representations.



http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.422.8795&rep=rep1&type=pdf

1 Velocity scooter (without controls)

uniform varying motion

v(t) in m/s

10 9 8 7 6 5 4 3 2 1

0 1 2 3 4 5 6 7 8 9

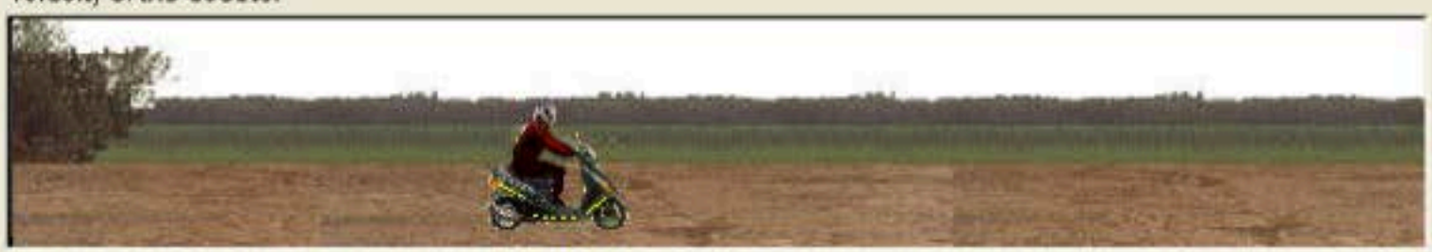
t in sec

scooter starting from stand still

initial velocity v(0) 0 m/s

acceleration a 0.50 m/s²

velocity of the scooter



simulation delay

velocity meter 0 - 15 m/s

Assignment Image

Close

You have 1 attempt left

Assignment

Try to make the velocity of the scooter 8 m/s in 4 seconds.

You've to do this by giving the right value to the acceleration.

You may try two times.

Start Close

DEMIST - tsp.mer

File Edit View Controller Window Help

ERs

Graph: P and N v Time

N (Field 1)
P (Field 1)

Animation: Change(X) and X

Field 1

dN/dT : 0.3764
N : 23.8403
dP/dT : 0.1076
P : 7.4870

Table: N and P

	N	P
0	20.0	5.0
5	30.4	5.4
10	37.1	6.6
15	34.9	8.5
20	26.6	10.2
25	18.6	10.8
30	13.6	10.5
35	11.3	9.7
40	10.9	8.8
45	12.0	7.9
50	14.4	7.4
55	17.9	7.1
60	21.6	7.2
65		
70		
75		

Simple Equation

Field 1

Simple Mathematical Model

dN/dT = Potential(N) - Killed(N)

0.38 5.73 5.35

dP/dT = Potential(P) - Dead(P)

0.11 0.48 0.37

NB if P density was 0, killed (N) would be 0 and the behaviour would be the same as single species limited growth.

Chart: b(N) and d(P)

b (N) d (P)

Field 1

Phaseplot: P v N

Field 1

P

N

Animated N and P

Field 1

Controller

Run Start

Review

Task Initialise Next Unit

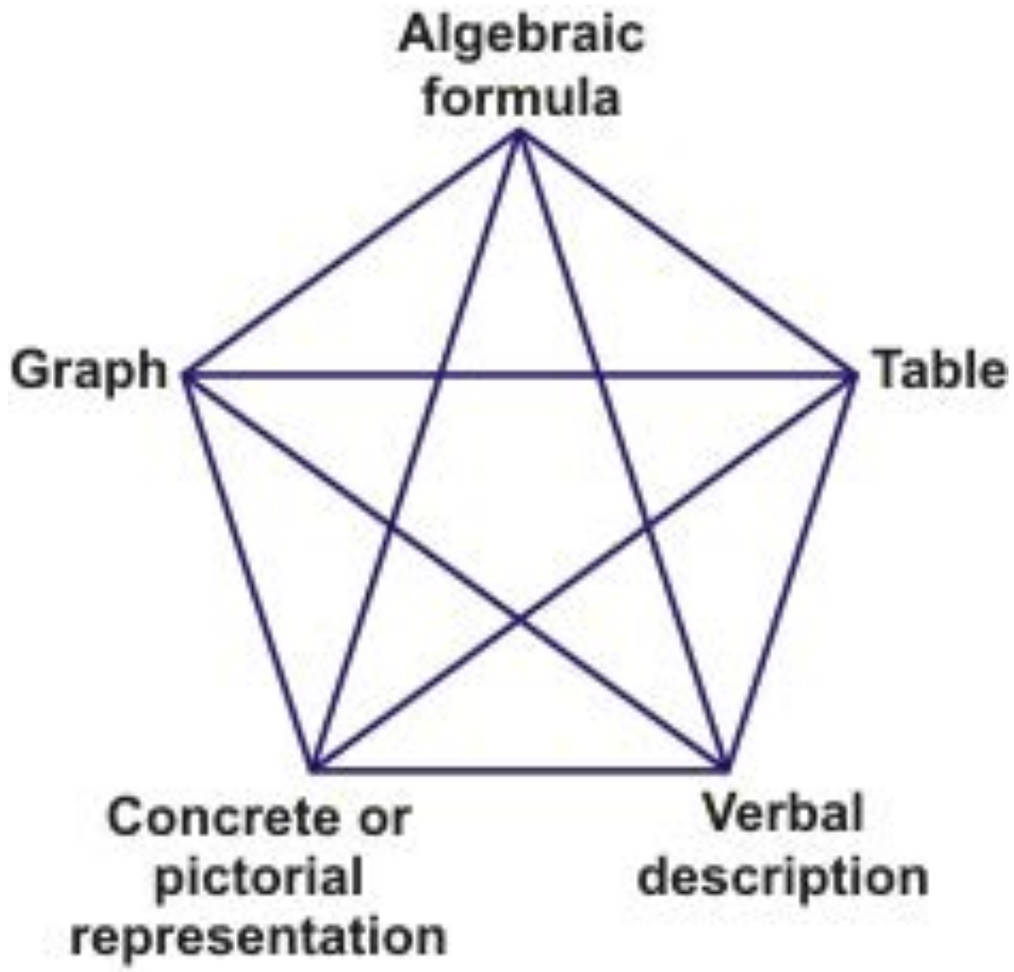
For Help, press F1

Learning mode

Why is an activity a_1 useful for an activity a_2?

Shift Representation:

a_2 represents contents differently from a_1



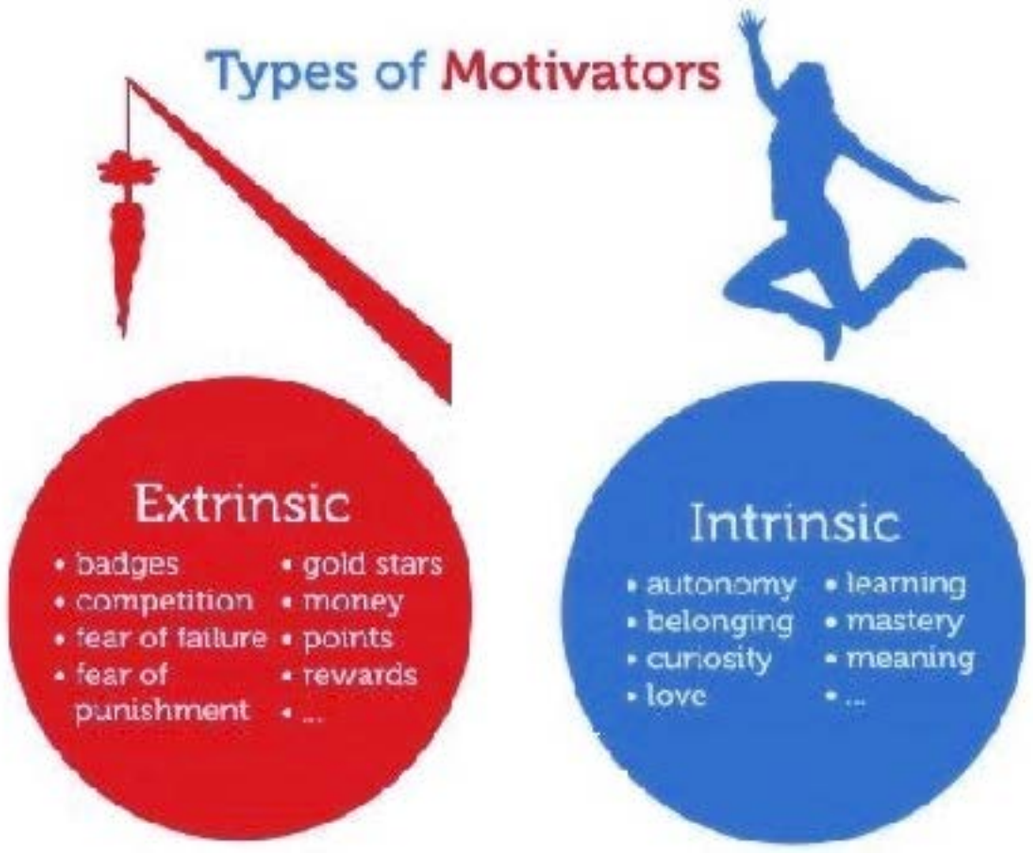
http://www.nctm.org/classroom-resources/lessons/Graphical-Representations-for-the-Number-of-Hits/

Why is an activity a_i useful for an activity a_{i+1}?

1. a_i is prerequisite to a_{i+1}
2. a_i is pre-activates cognitive schema for a_{i+1}
3. a_{i+1} is varies the representations used in a_1
4. ...

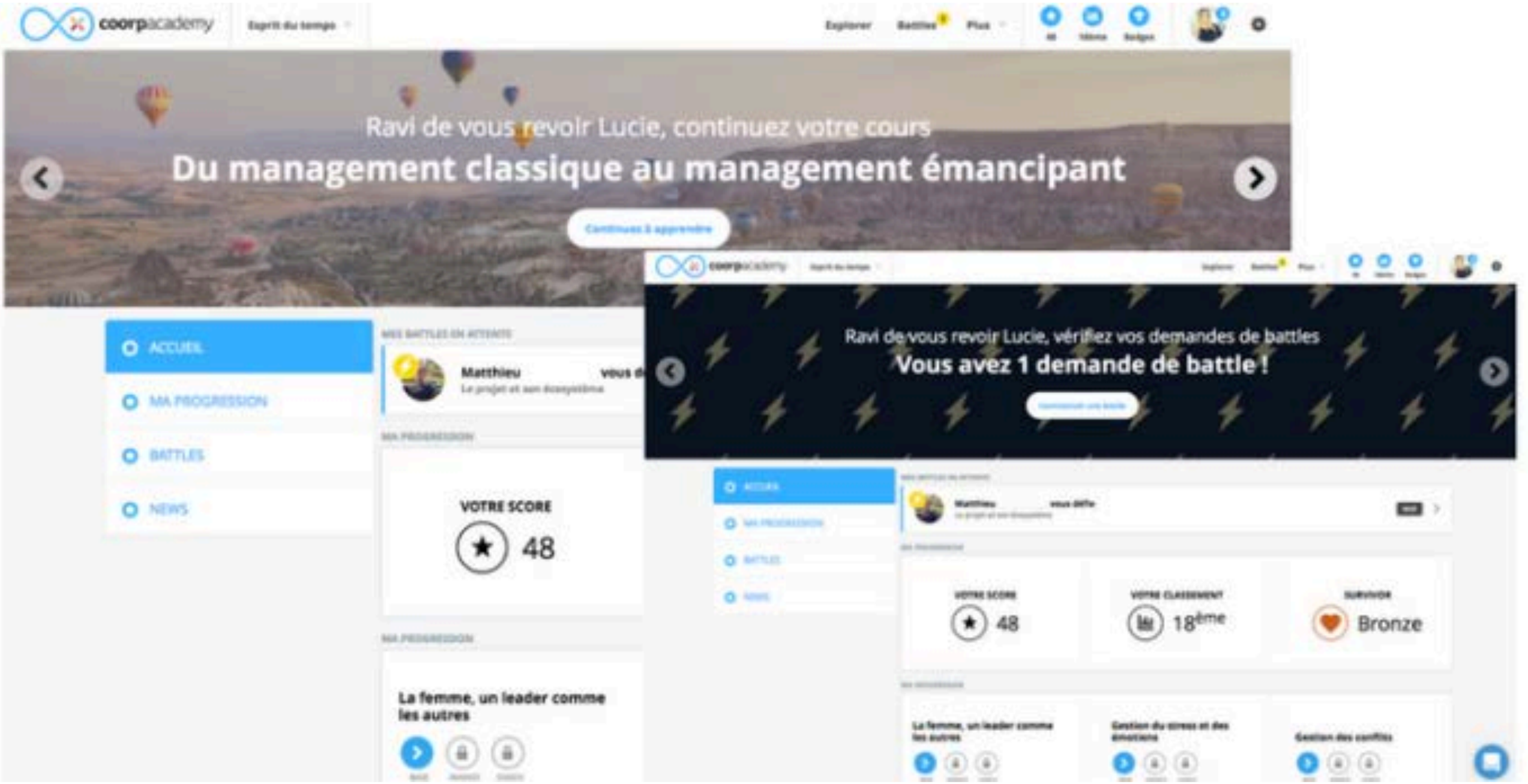
Why did you take this class ?

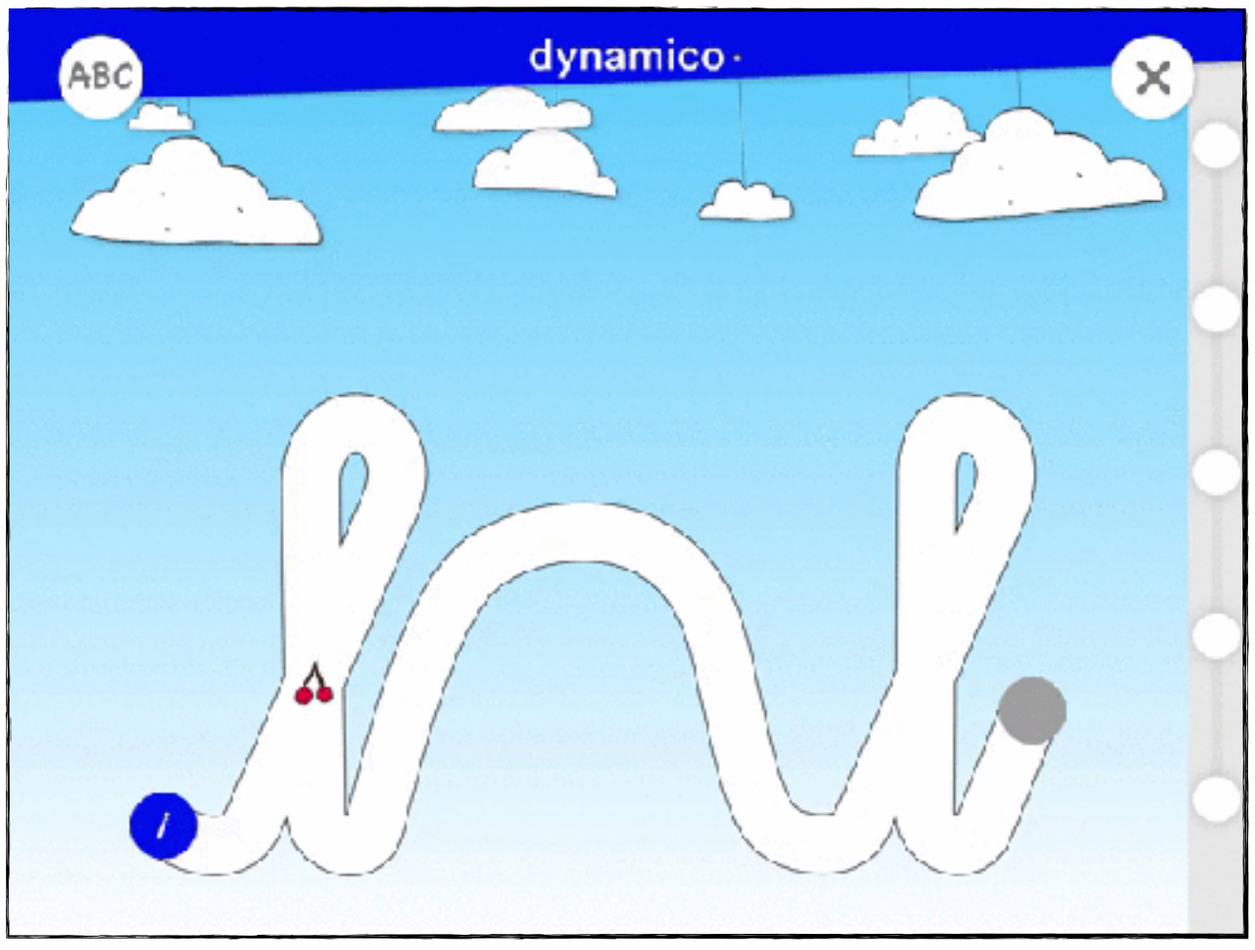
A. I am passioned about education
B. I have been told it's an easy one
C. I had nothing on Tuesday morning
D. I would like to create an EdTech start-up

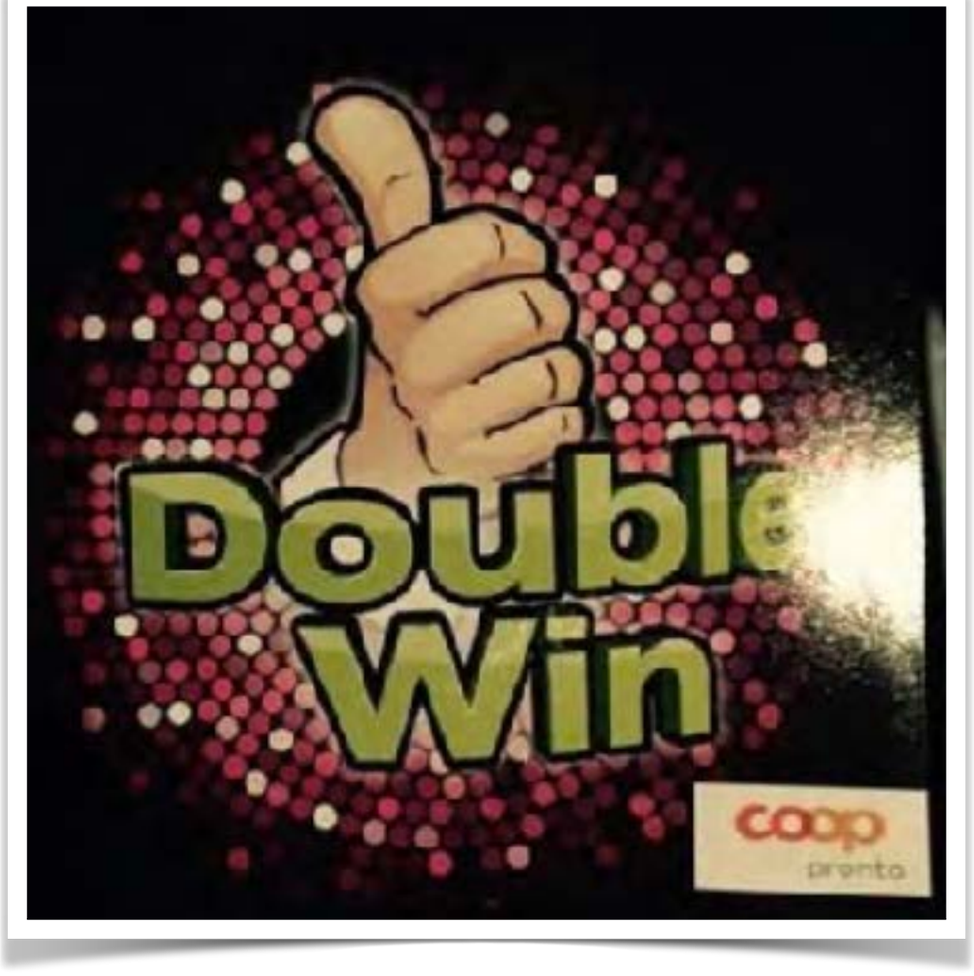


Gamification

(even in corporate training)







Game ≠ Gamification

Gamification

Progressivity

easier

Inverted Progressivity

motivating

Why is an activity a_1 useful for an activity a_2?

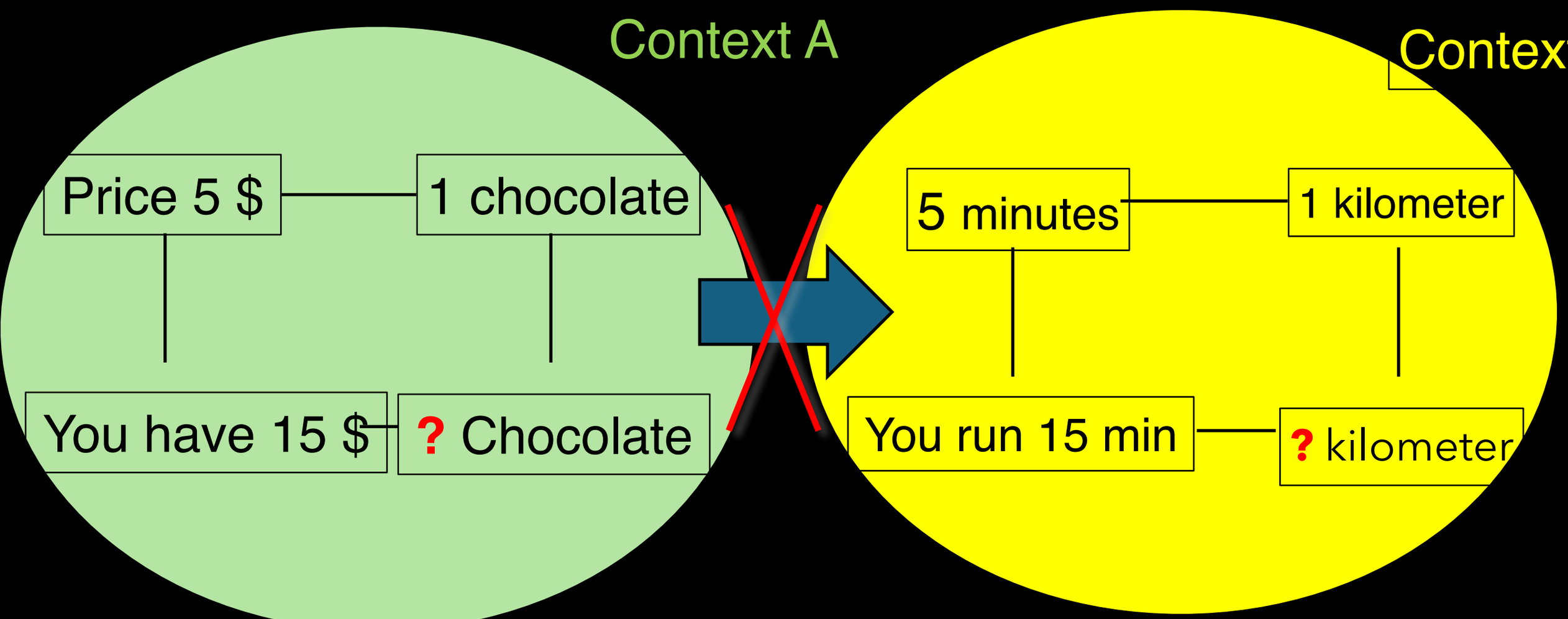
Motivation:

a_1 frustrate learners so that they want to do a_2

« what you learned before enabled you to solve problems so far, but here are new problems »

Why is an activity a_i useful for an activity a_{i+1}?

1. a_i is prerequisite to a_{i+1}
2. a_i is pre-activates cognitive schema for a_{i+1}
3. a_{i+1} is varies the representations used in a_1
4. a_i brings motivation for a_{i+1}
5. ...



Transfering knowledge between contexts
is not (often) **spontaneous**

Why is an activity a_1 useful for an activity a_2?

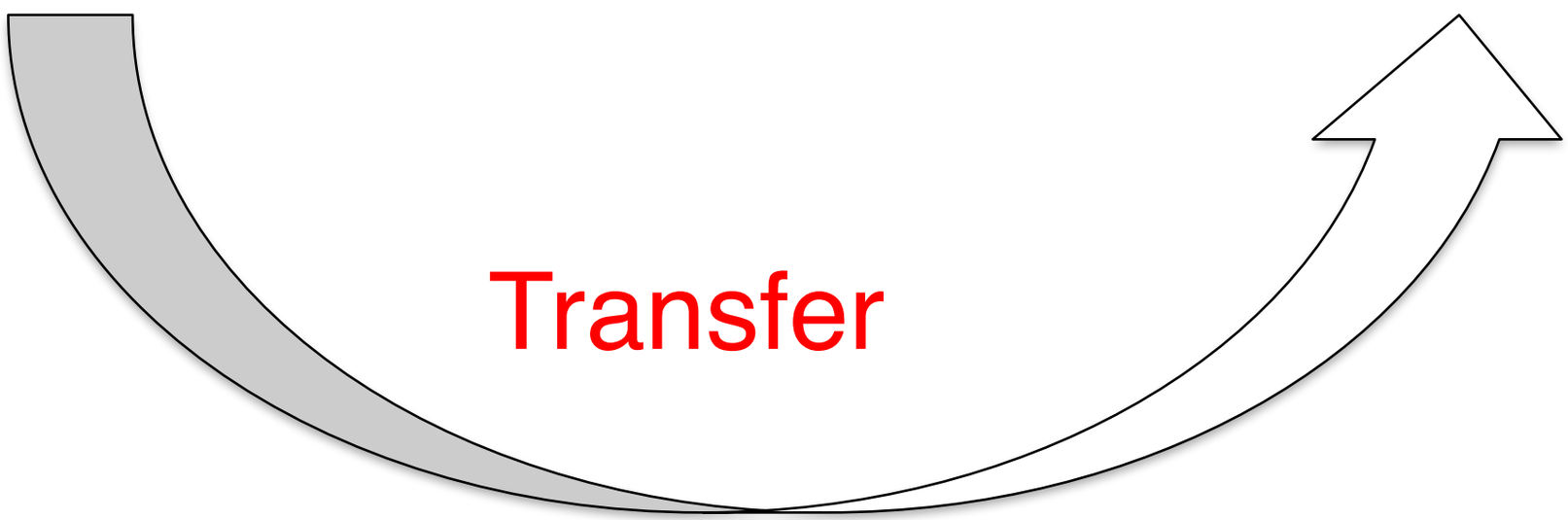
$$5\,X < 27$$

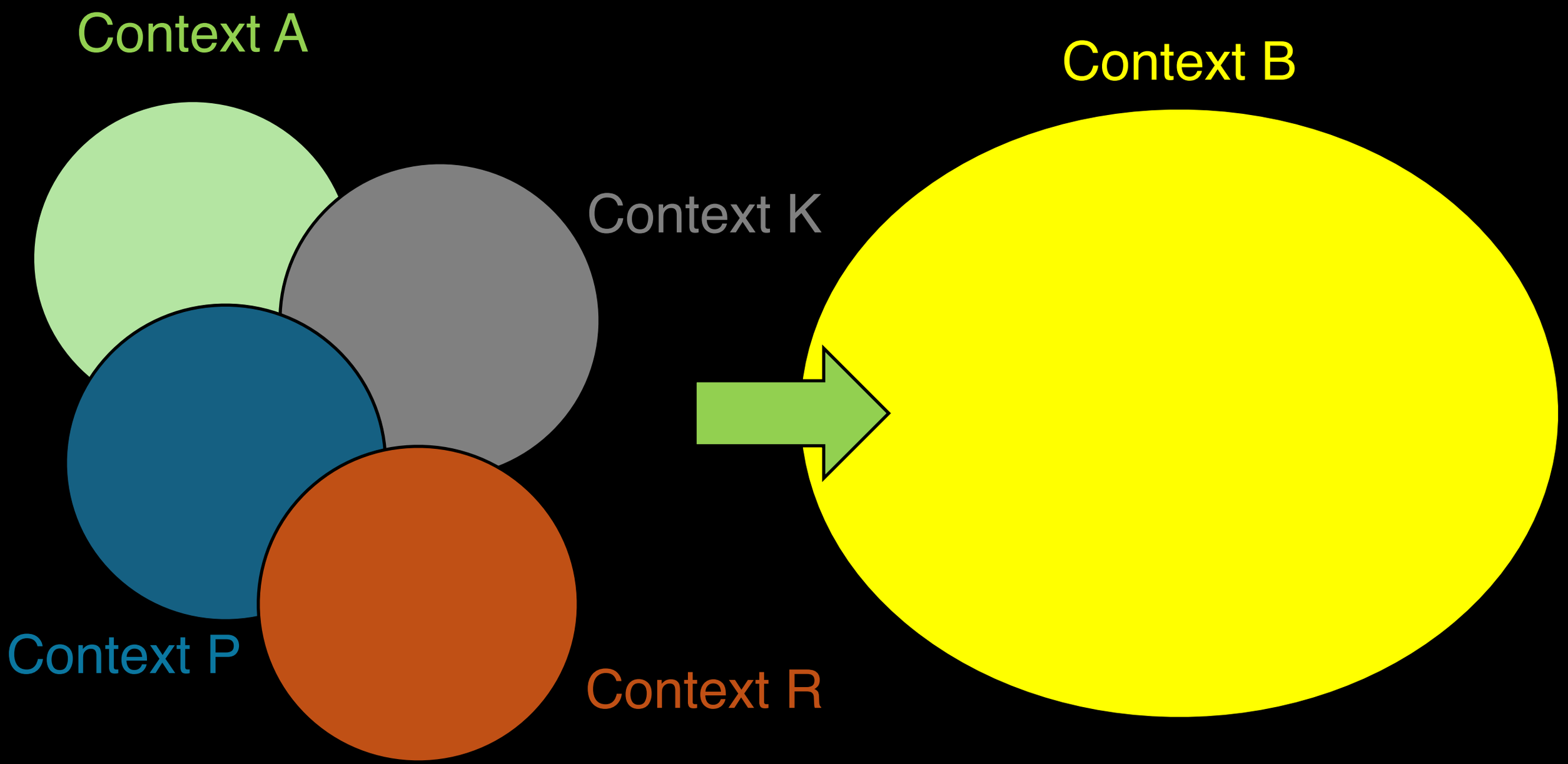
a_i

A chocolate bar costs 5 CHF.
How many bars can you buy for 25 CHF

a_i

A man walks 1 km in 5 min.
How many km can he walk in 25 min





Transfer can be trained by varying contexts

Why is an activity a_i useful for an activity a_{i+1}?

1. a_i is prerequisite to a_{i+1}
2. a_i is pre-activates cognitive schema for a_{i+1}
3. a_{i+1} is varies the representations used in a_1
4. a_i brings motivation for a_{i+1}
5. a_{i+1} practices the transfer of skills acquired in a_1
6.

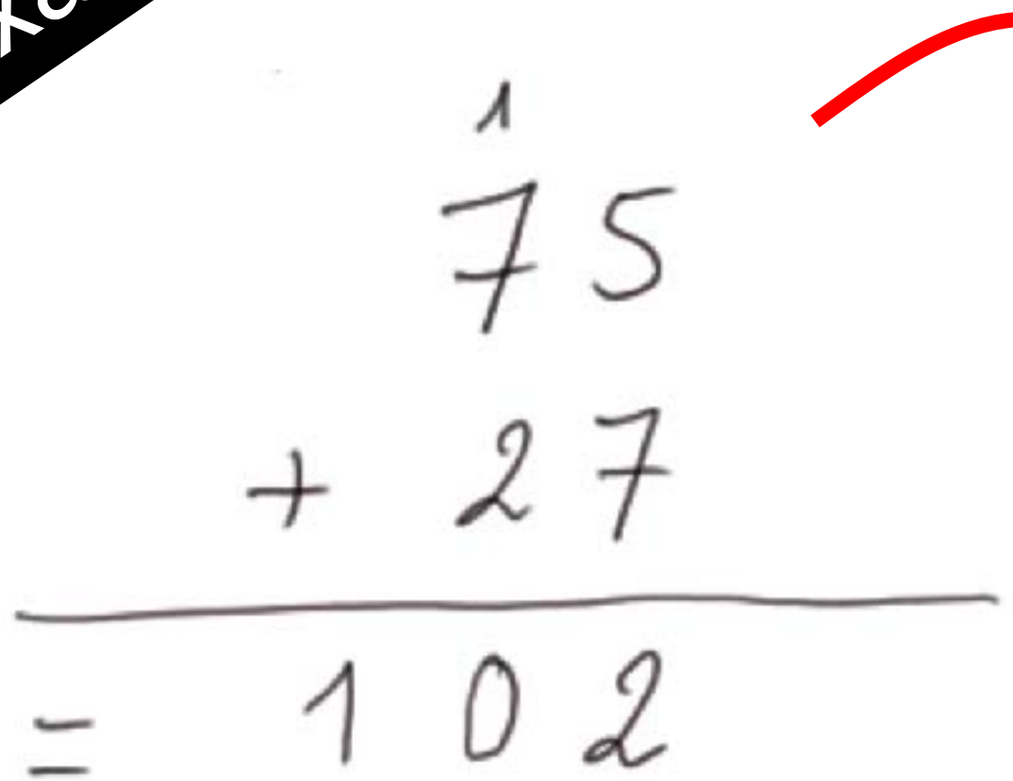
How to favor skills transfer ?

1. By reinstantiating skills in various contexts

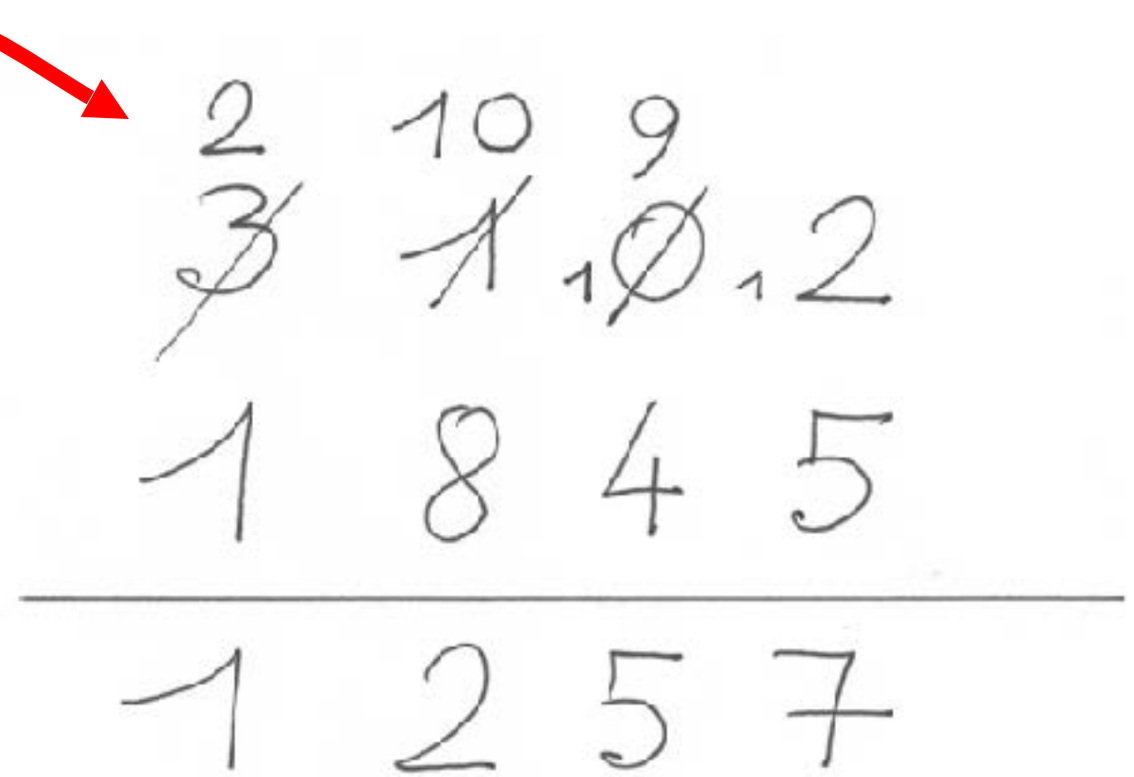
2. By targetting a deeper understanding

Example 1

How to foster skills transfer ?



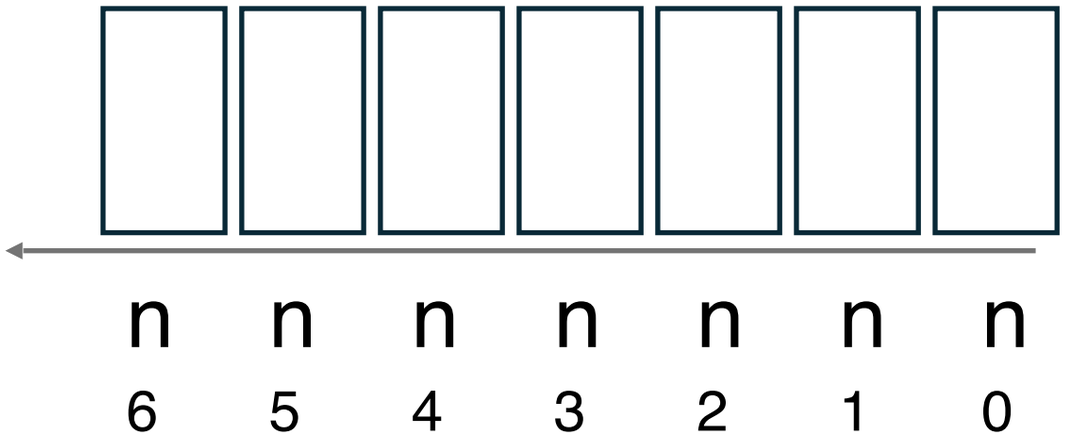
Written Addition



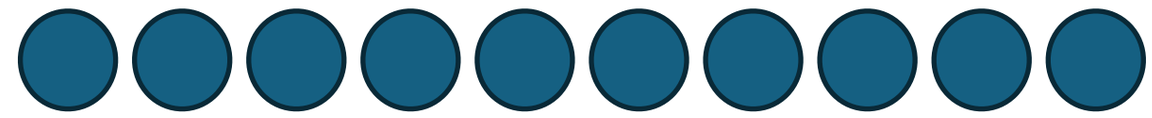
Written Substraction

Why do we write numbers the way we do

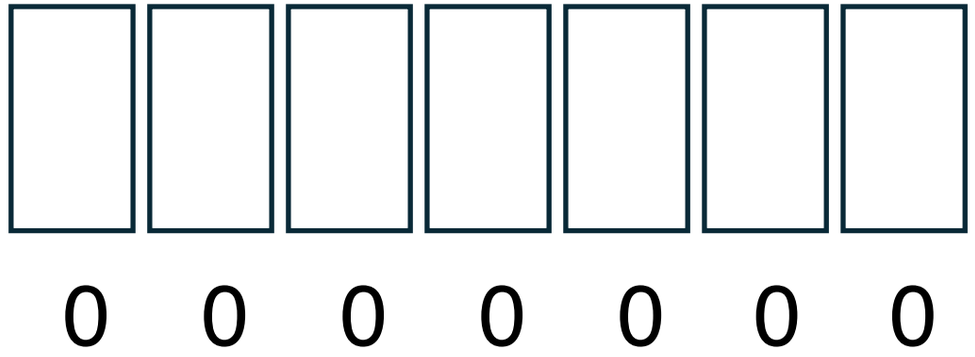
How does a positional numeral system work?



Base n

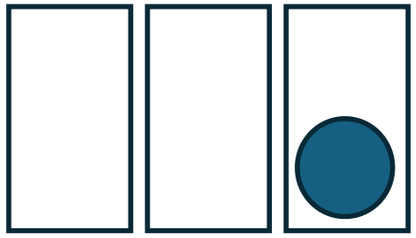


10 tokens

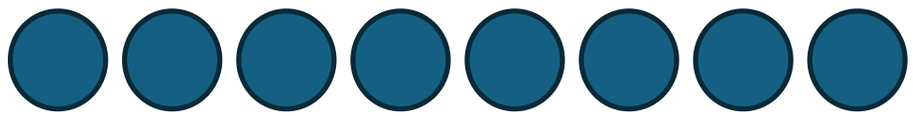


Base 3

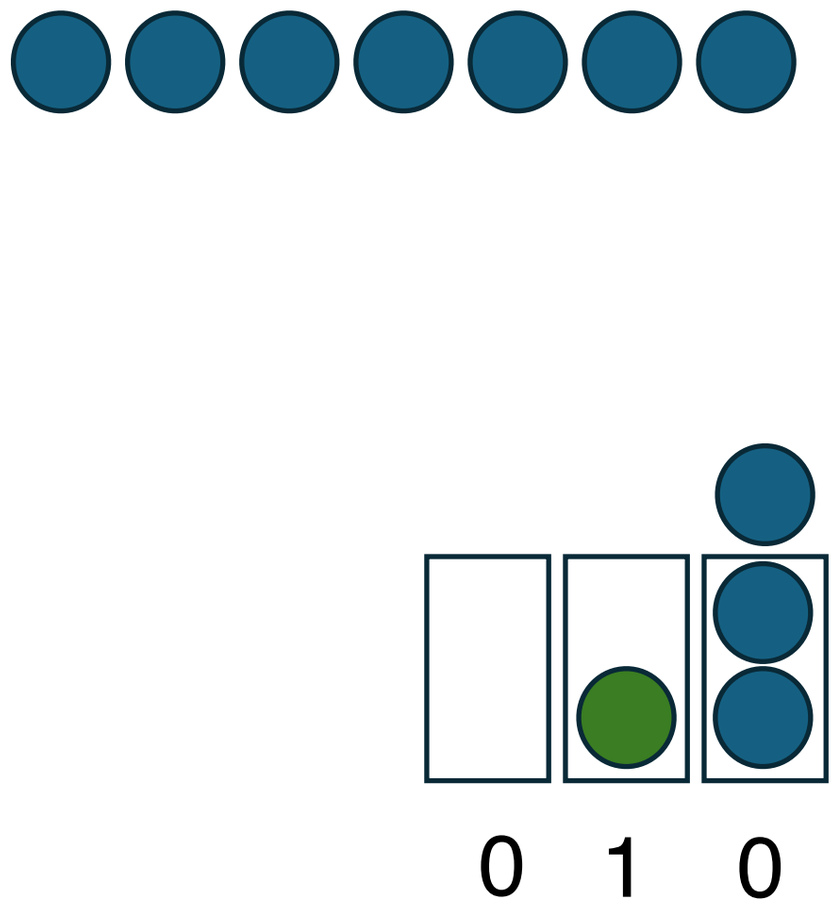
How to write 10 in base 3 ?

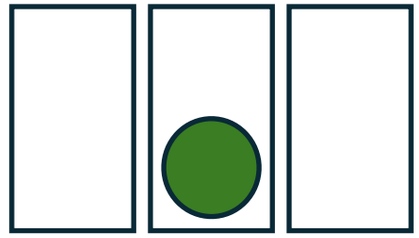


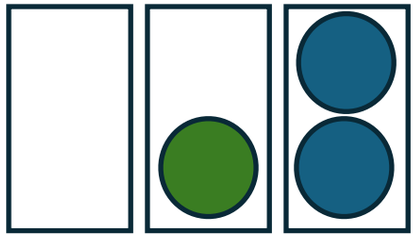
0 0 1

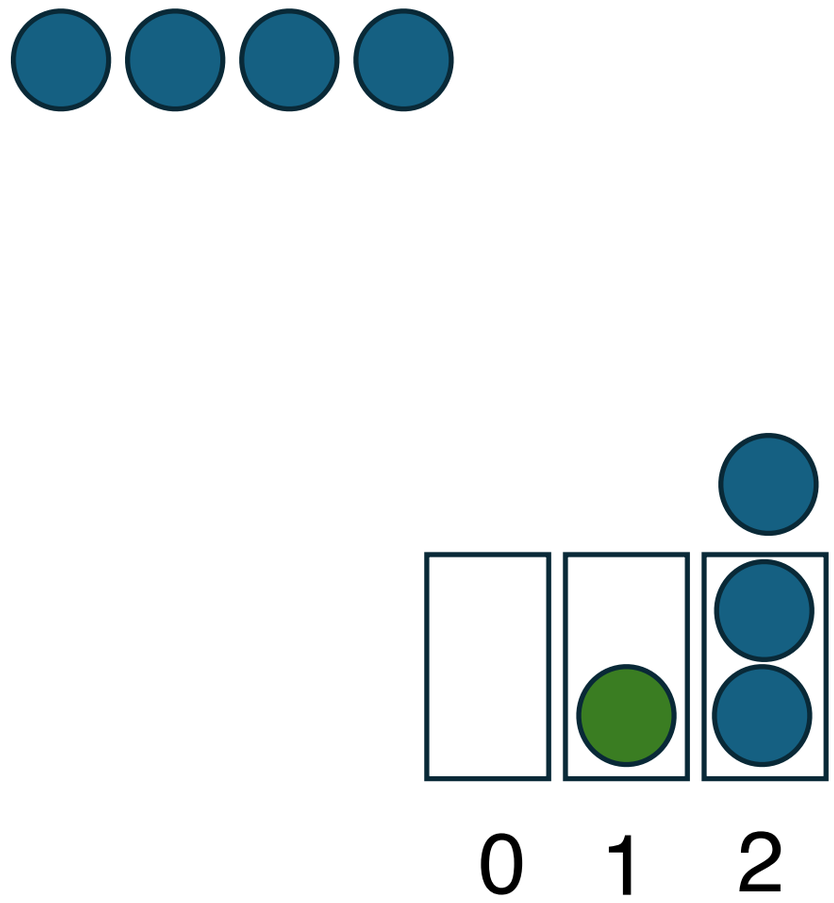


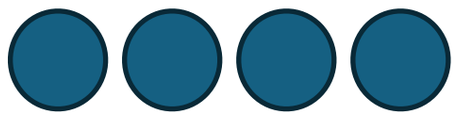
0 0 2

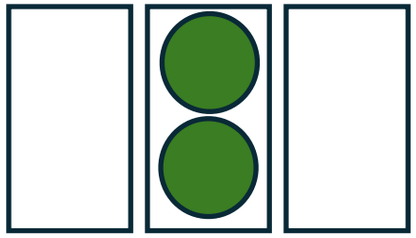

0 1 0


0 1 2

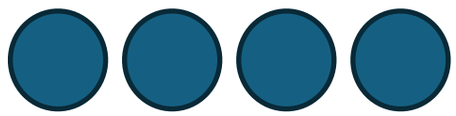

0
1
2

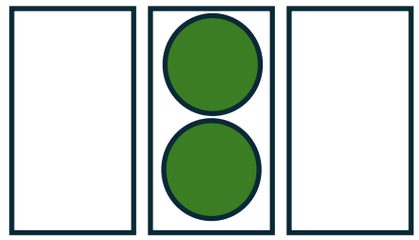

0 1 2



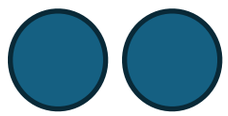


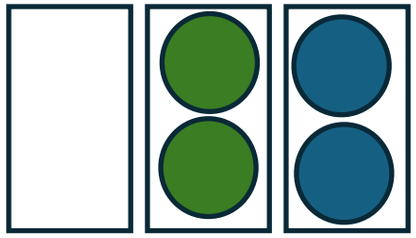
0 2 0



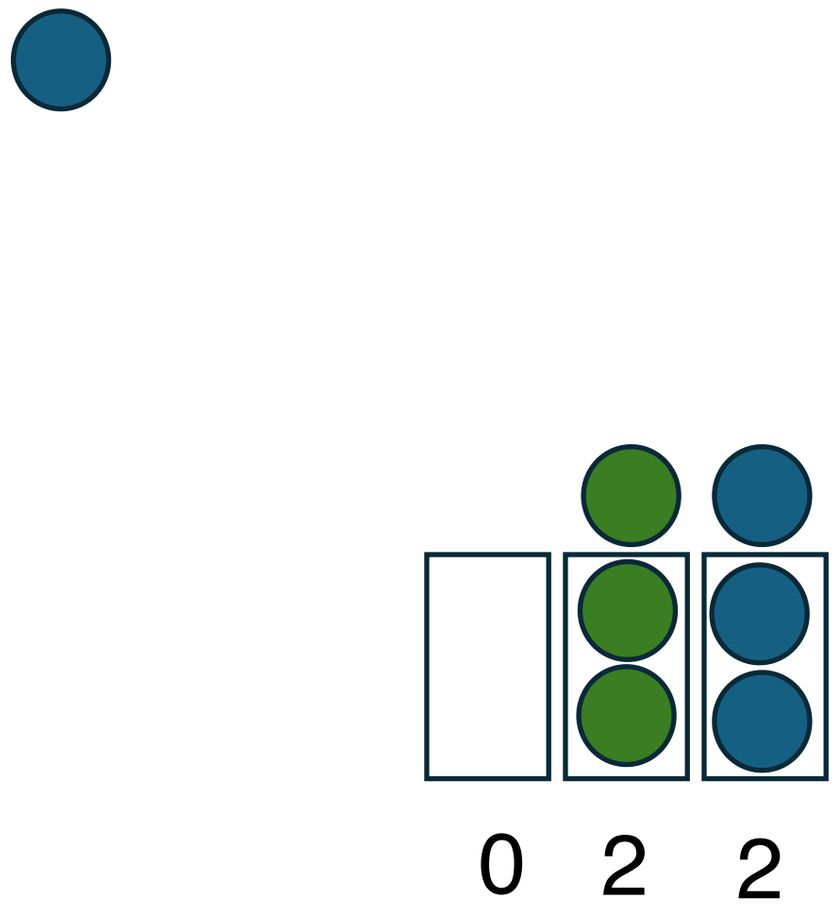


0 2 1

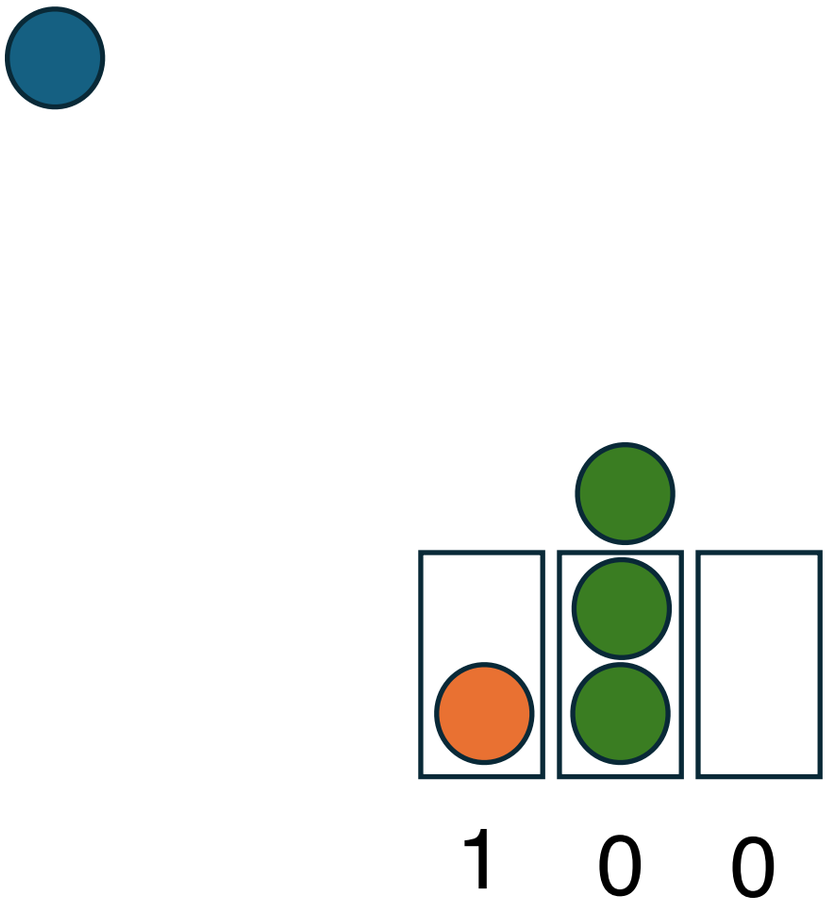


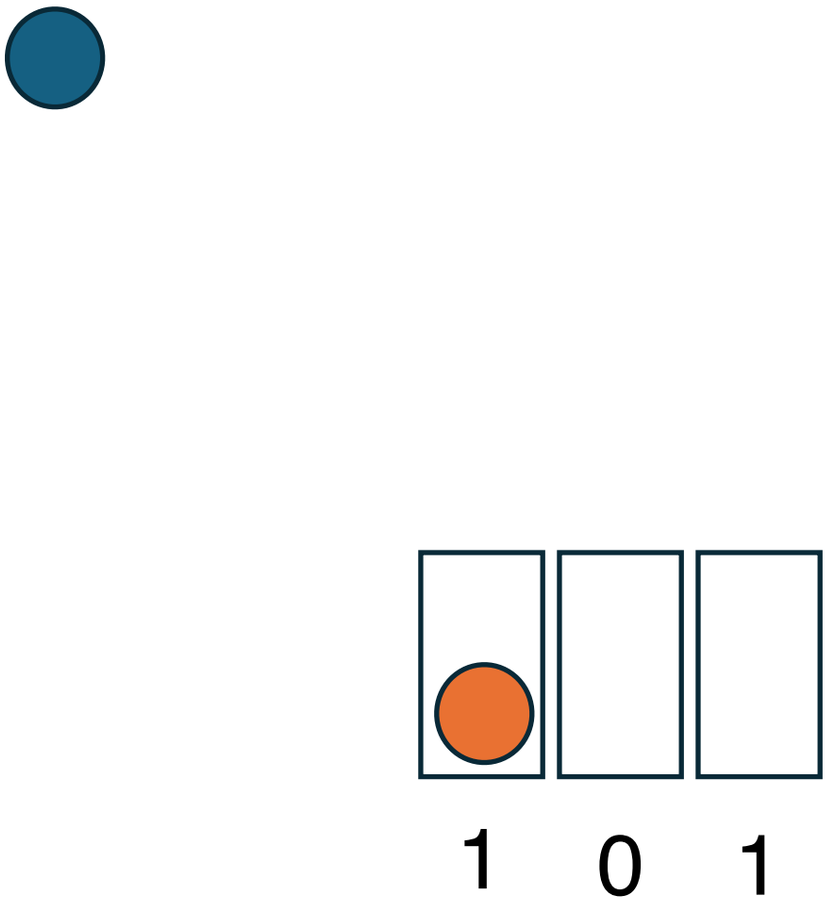


0 2 2



0 2 2


1 0 0

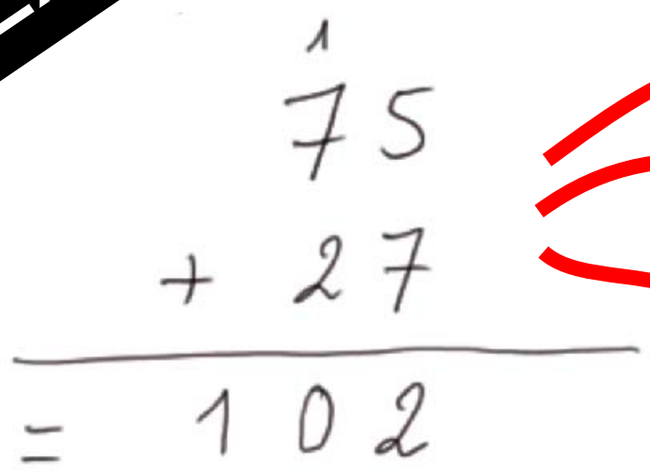


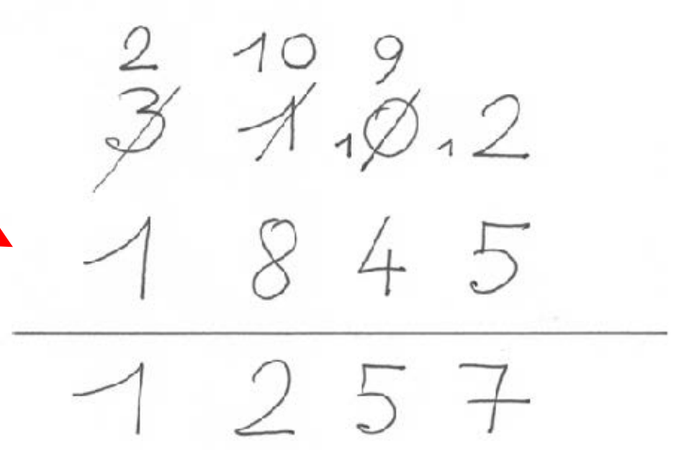
10 (base 10) = 101 (base 3)

Example 1

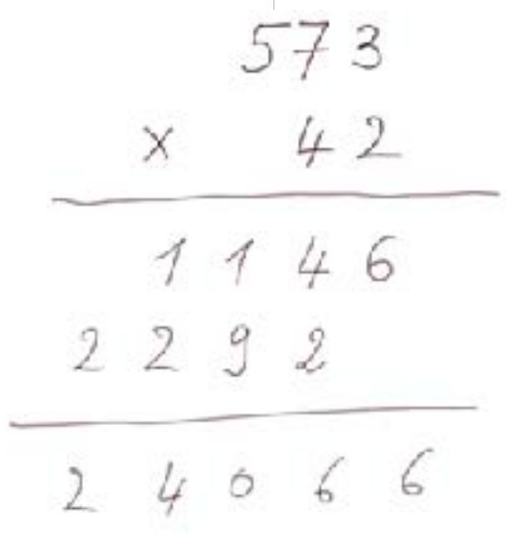
How to foster skills transfe

Written
Addition

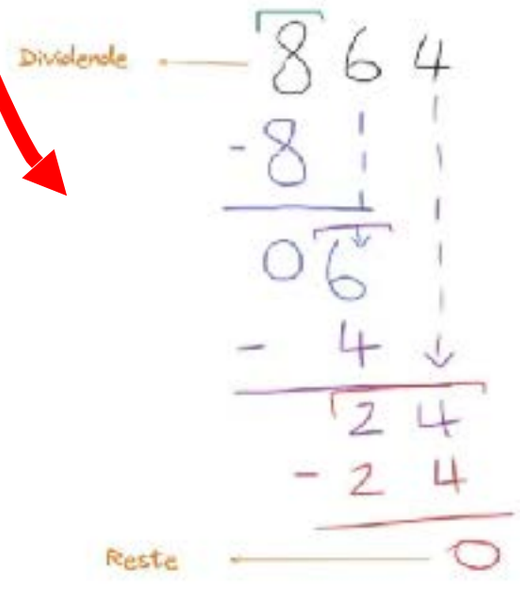


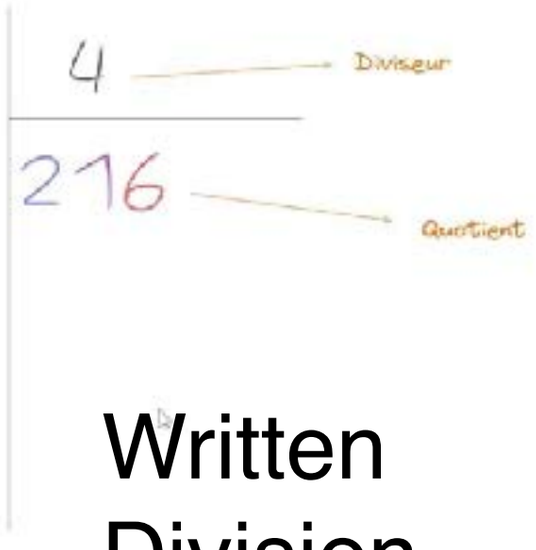


Written
Substraction

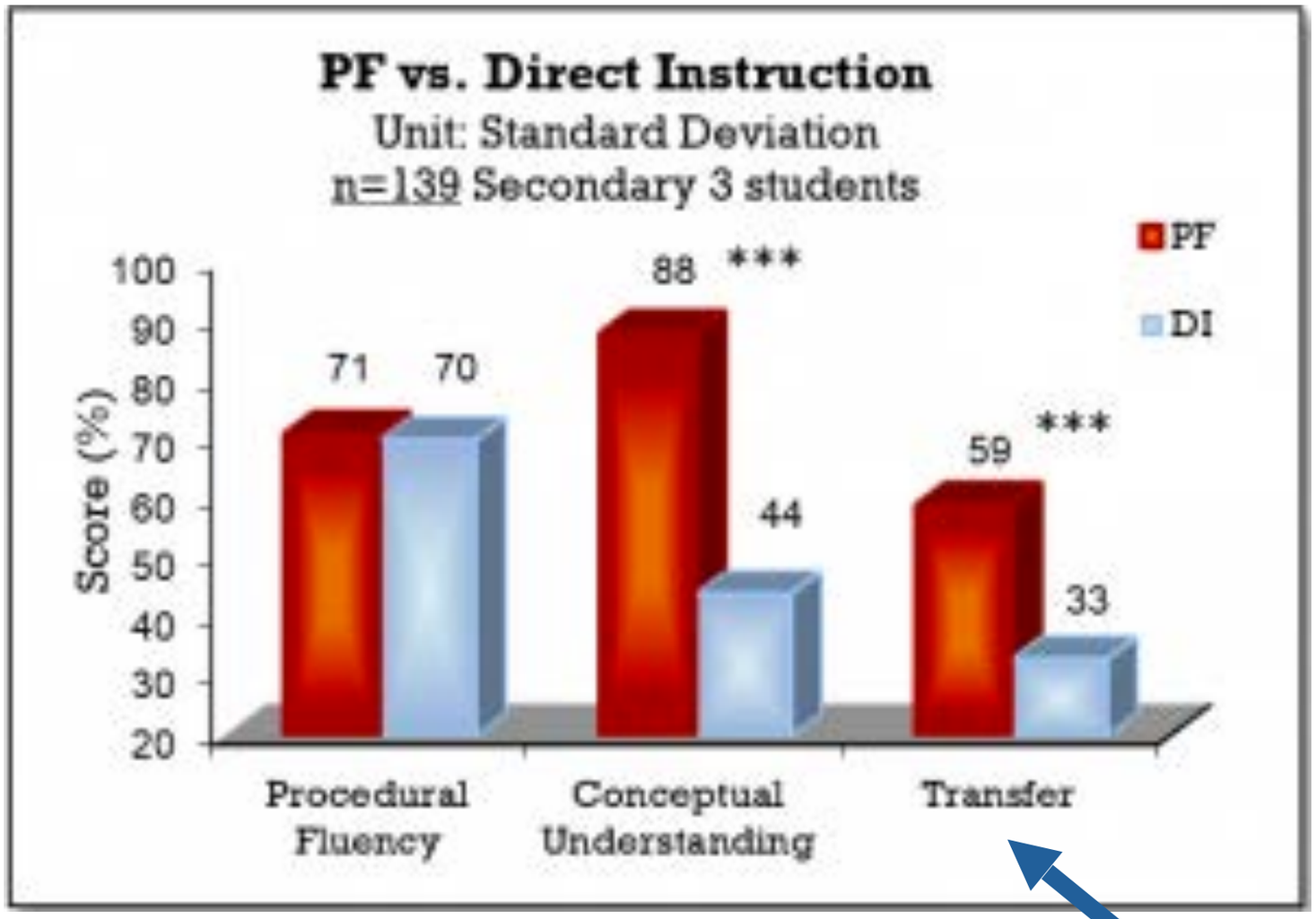


Written
Multiplication





Written
Division



Kapur, M., & Bielaczyc, K. (2012). Designing for productive failure. *Journal of the Learning Sciences*, *21*(1), 45-83.

There is no magic wand in education

166 experimental comparisons:

- Average effect size, Hedge's *g* = 0.36
- With fidelity to PF, Hedge's *g* = 0.58

…ma & Kapur, 2021

Why is an activity a_i useful for an activity a_{i+1}?

1. a_i is prerequisite to a_{i+1}
2. a_i is pre-activates cognitive schema for a_{i+1}
3. a_{i+1} is varies the representations used in a_1
4. a_i brings motivation for a_{i+1}
5. a_{i+1} practices the transfer of skills acquired in a_1
6. a_{j+1} summarizes in $\{a_1, \ldots, a_j\}$
7. a_{i+2} contrasts examples used in a_1 and a_2

How do people learn ?

☑ by exploration, trial and error

✓ **by incremental mastery**

- by verbal elaboration

There is no « modern » pedagogy