

# CS-411 : Digital Education & Learning Analytics

## Chapter 3: Mastery learning

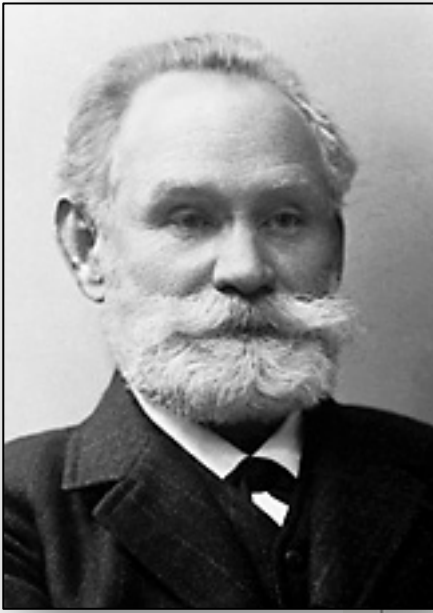
1. Sensory Memory
2. Working Memory
3. Articulatory loop
4. Verbal/Phonological Loop
5. Articulatory suppression
6. Visuo-spatial sketchpad
7. Long-Term Memory
8. Cognitive Load
9. Intrinsic Cognitive Load
10. Extrinsic Cognitive Load
11. Germane Cognitive Load
12. Cognitive Overload
13. Split-Attention Effect
14. Induction
15. Deduction
16. Analogy
17. Fact
18. Class
19. Procedure
20. Law
21. System
22. Knowledge Taxonomy
23. Cognitive tasks taxonomy
24. Bloom's taxonomy
25. D'Hainaut's taxonomy

LAST  
WEEK

26. Pedagogical Objectives
27. Learning Outcomes
28. Declarative knowledge
29. Procedural knowledge
30. Elicitation
31. Proceduralisation
32. Compilation
33. Free recall
34. Recognition
35. Imitation
36. Conceptualisation
37. Application
38. Exploration
39. Mobilisation
40. Problem solving
41. Transversal skills
42. Heuristic Knowledge
43. Formal education
44. No-formal Education
45. Informal Education
46. Metacognition
47. Regulation
48. ~~Pre-requisites~~
49. ~~Pre-representations~~
50. ~~Instructional Engineering~~

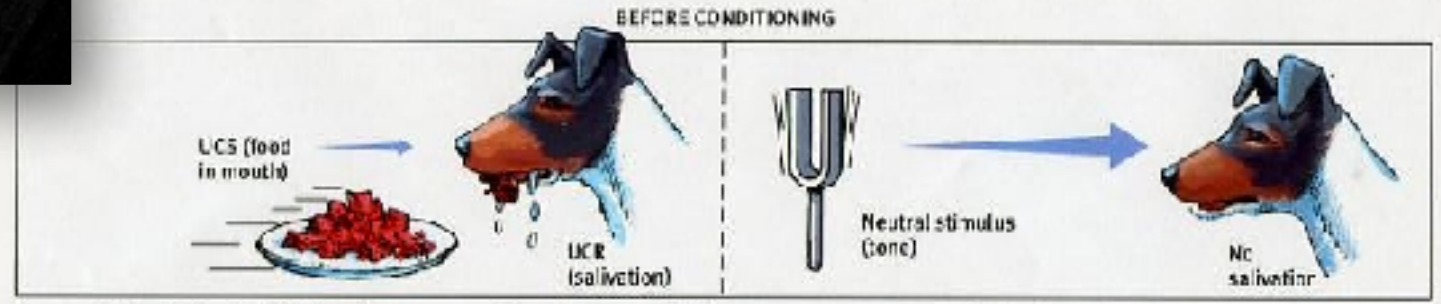
dogs

How do ~~people~~ learn ?



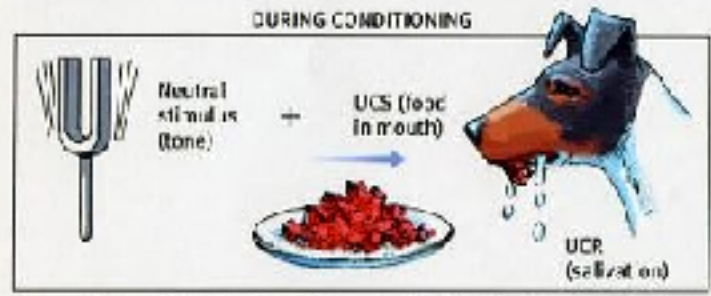
Yvan Pavlov, 1849-1936

# Classical Conditioning

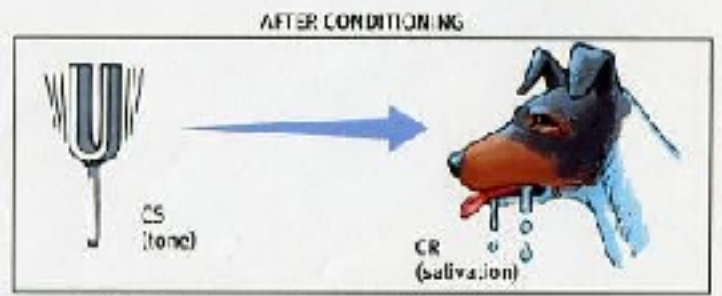


An unconditioned stimulus (UCS) produces an unconditioned response (UCR).

A neutral stimulus produces no salivation response.



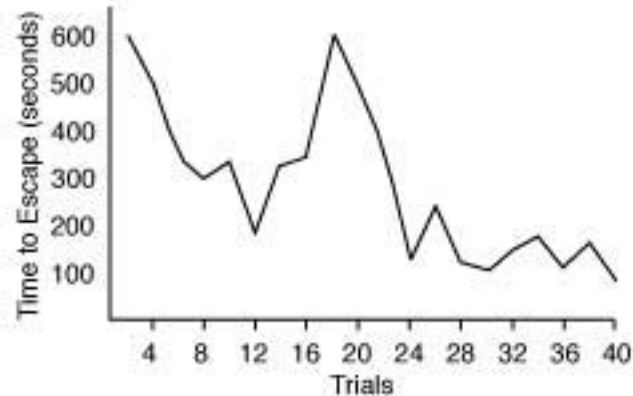
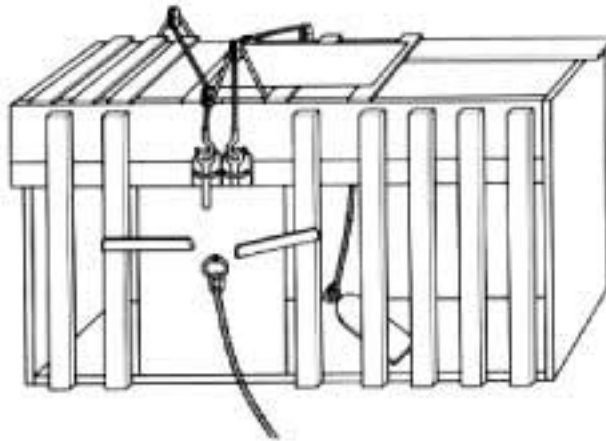
The unconditioned stimulus is repeatedly presented just after the neutral stimulus. The unconditioned stimulus continues to produce an unconditioned response.



The neutral stimulus alone now produces a conditioned response (CR), thereby becoming a conditioned stimulus (CS).



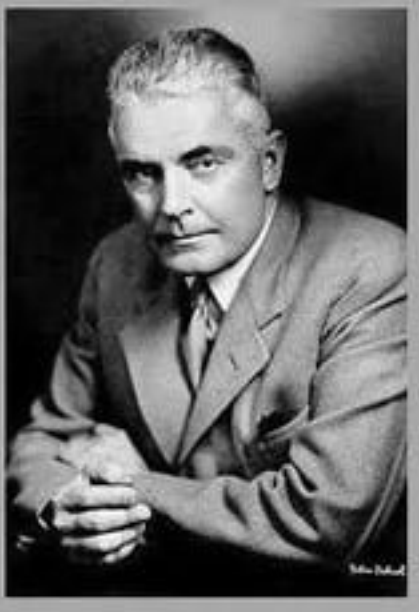
Edward L. Thorndike (1874 – 1949) [The Law of Effect](#) : any behavior that is followed by pleasant consequences is likely to be repeated, and any behavior followed by unpleasant consequences is likely to be stopped.



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

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

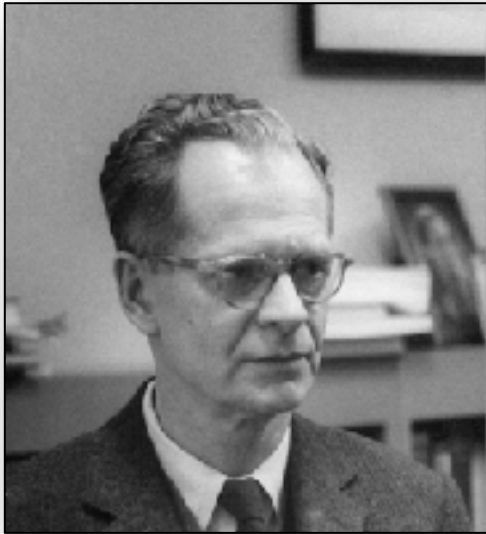
Edward L. Thorndike, The Law of Effect, The American Journal of Psychology  
Vol. 39, No. 1/4 (Dec., 1927), pp. 212-222: <http://www.jstor.org/stable/1415413>



## 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 beggarman and thief, regardless of his talents, penchants, tendencies, abilities, vocations, and race of his ancestors."

--John Watson, Behaviorism, 1930  
*The [Little Albert Experiment](#)*



Burrhus Frederic Skinner (1904-1990), [Operant Conditioning](#)

**SKINNER  
OPERANT  
CONDITIONING**

*"The Science of  
Changing  
Behavior"*

**BEHAVIOR**

**DESIRED  
POSITIVE**

**GIVE  
PLEASURE**

**REMOVE  
PAIN**

R  
E  
I  
N  
F  
O  
R  
C  
E  
R

**MAINTAIN  
POSITIVE**

**FADE  
Reinforcers**

**21  
DAYS**

**UNDESIRED  
NEGATIVE**

**GIVE  
PAIN**

**REMOVE  
PLEASURE**

R  
E  
I  
N  
F  
O  
R  
C  
E  
R

**"The closer you implement a Reinforcer to a Behavior, the greater the chance of changing behavior"**

# 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

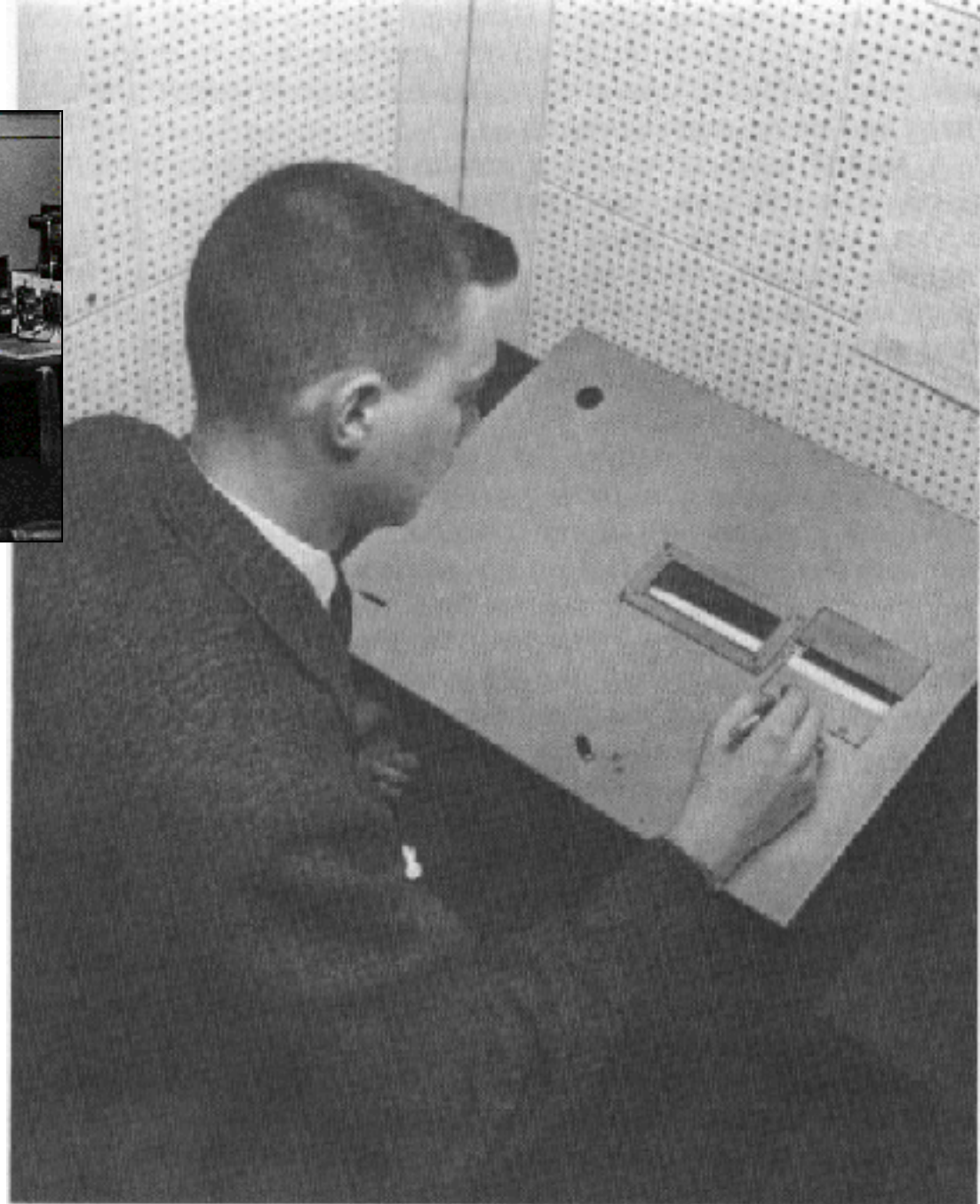
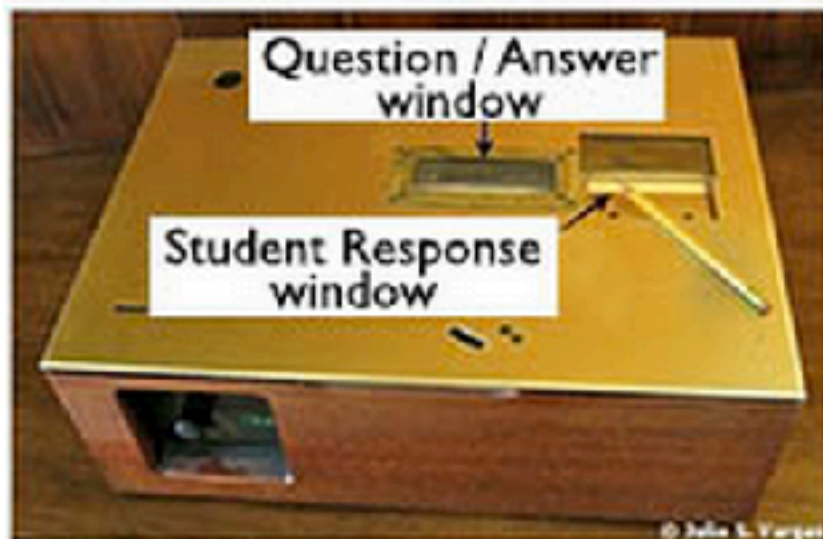
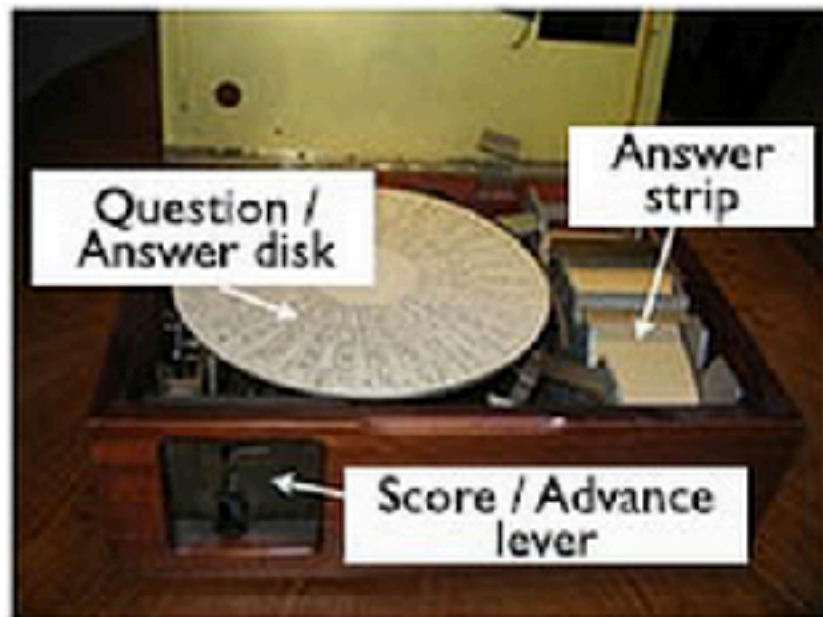


FIG. 11. Student at work in the self-instruction room. Material appears in the left-hand window. The student writes his response on a strip of paper exposed at the right.

## 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

# Frame-Based Models / e-learning

1. Information
2. Question
3. Feedback

1. Information
2. Question
3. Feedback

1. Information
2. Question
3. Feedback

1. Information
2. Question
3. Feedback

1. Information
2. Question
3. Feedback

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

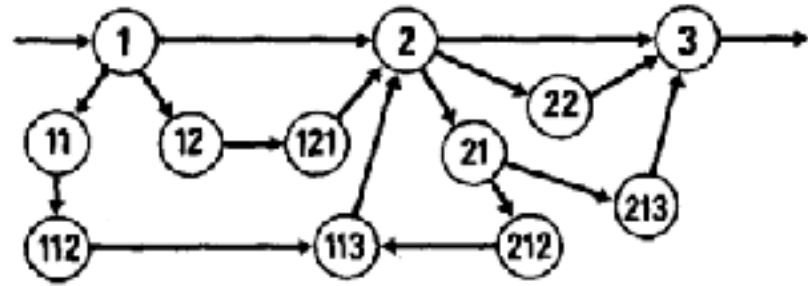
# Feedback

is the 1st principle

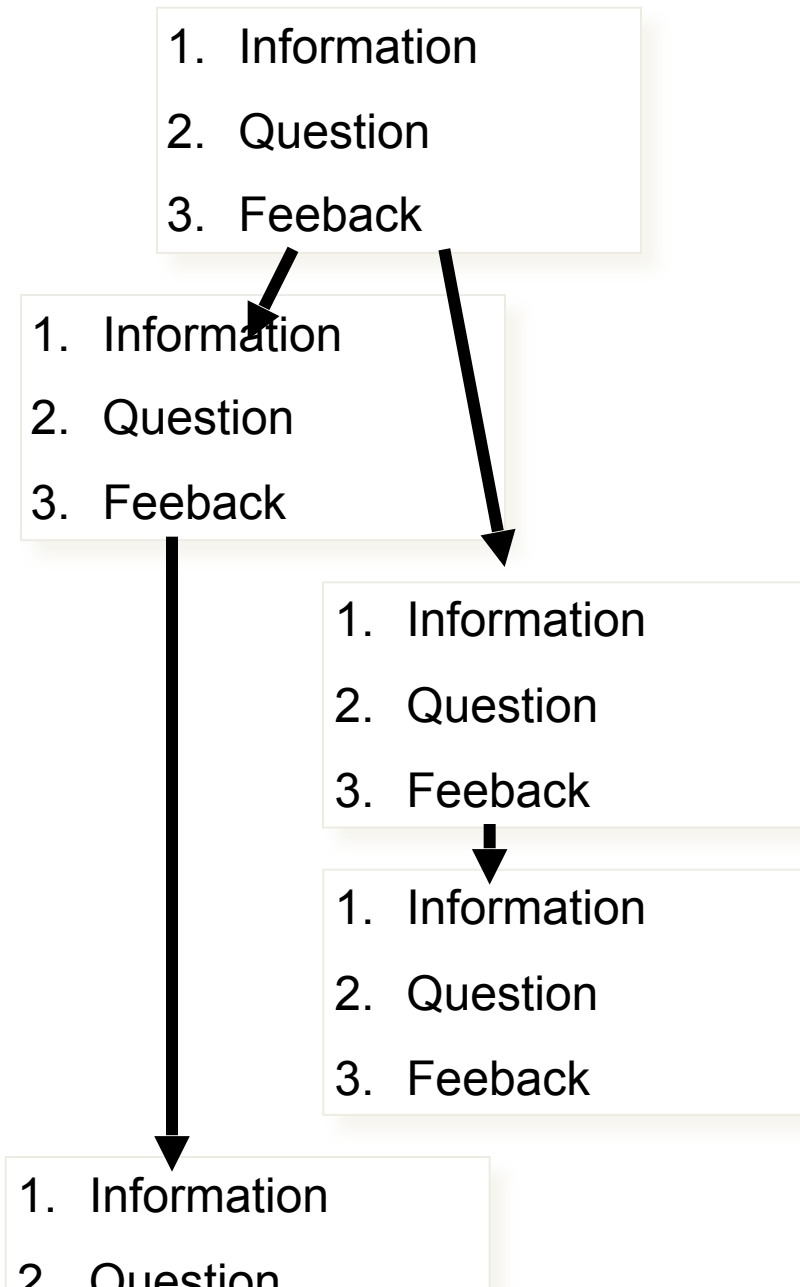
for pedagogical effectiveness



# Branched Instruction



AutoTutor, Crowder

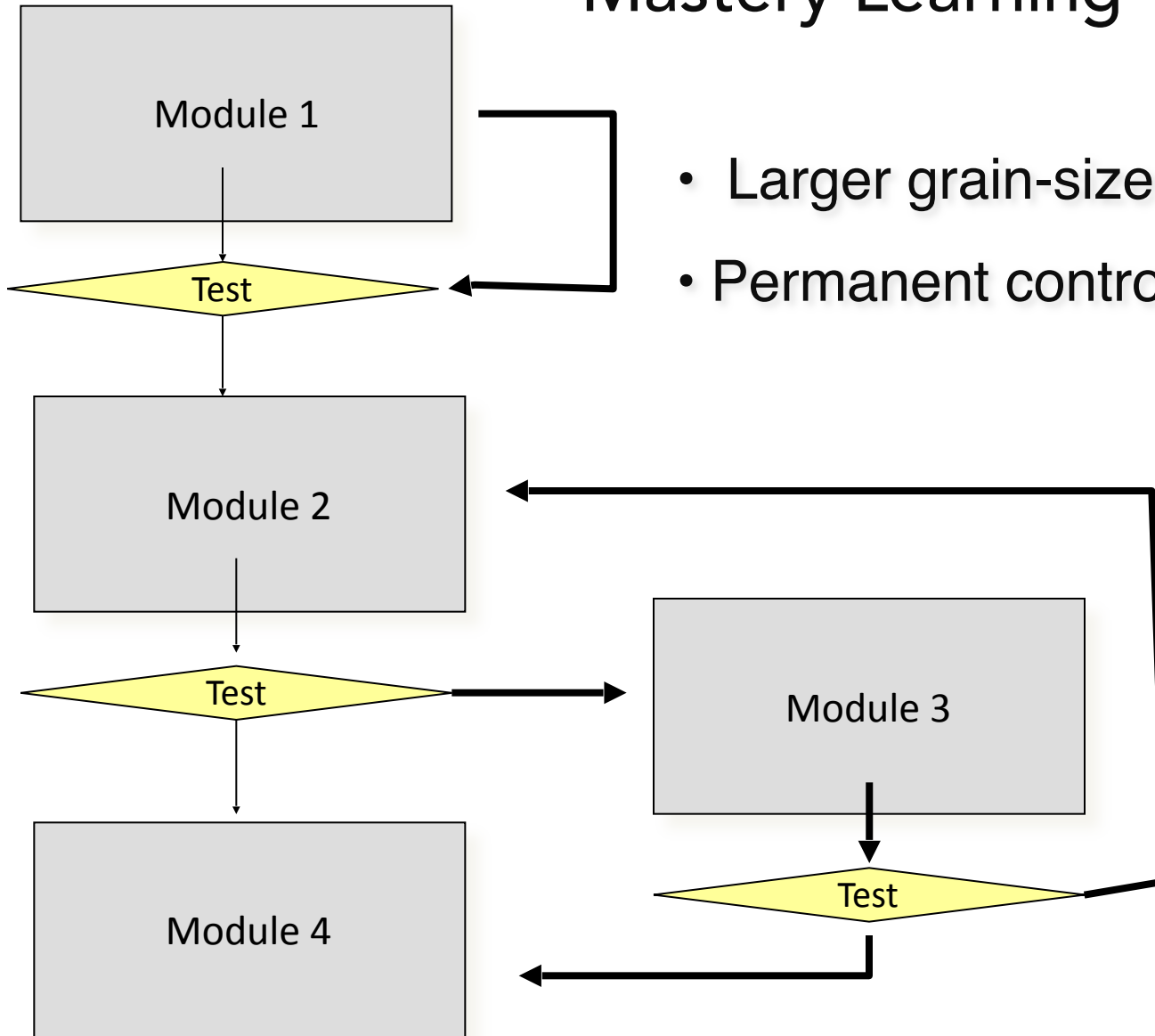


## Frame-Based Models

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

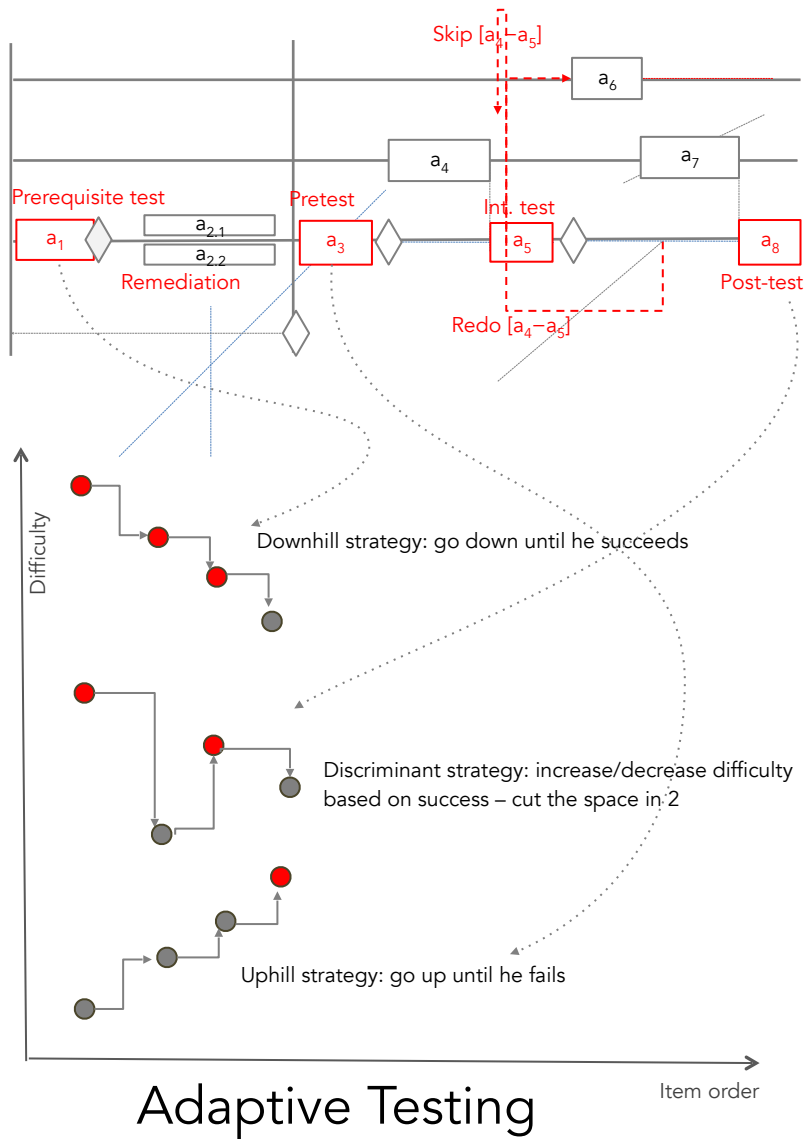
Based on behaviourism-inspired  
“programmed learning instruction”

# Mastery Learning



- Larger grain-size
- Permanent control of effectiveness





## Modular Instruction

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

**Pre-test:** Should the learner skip some modules ?

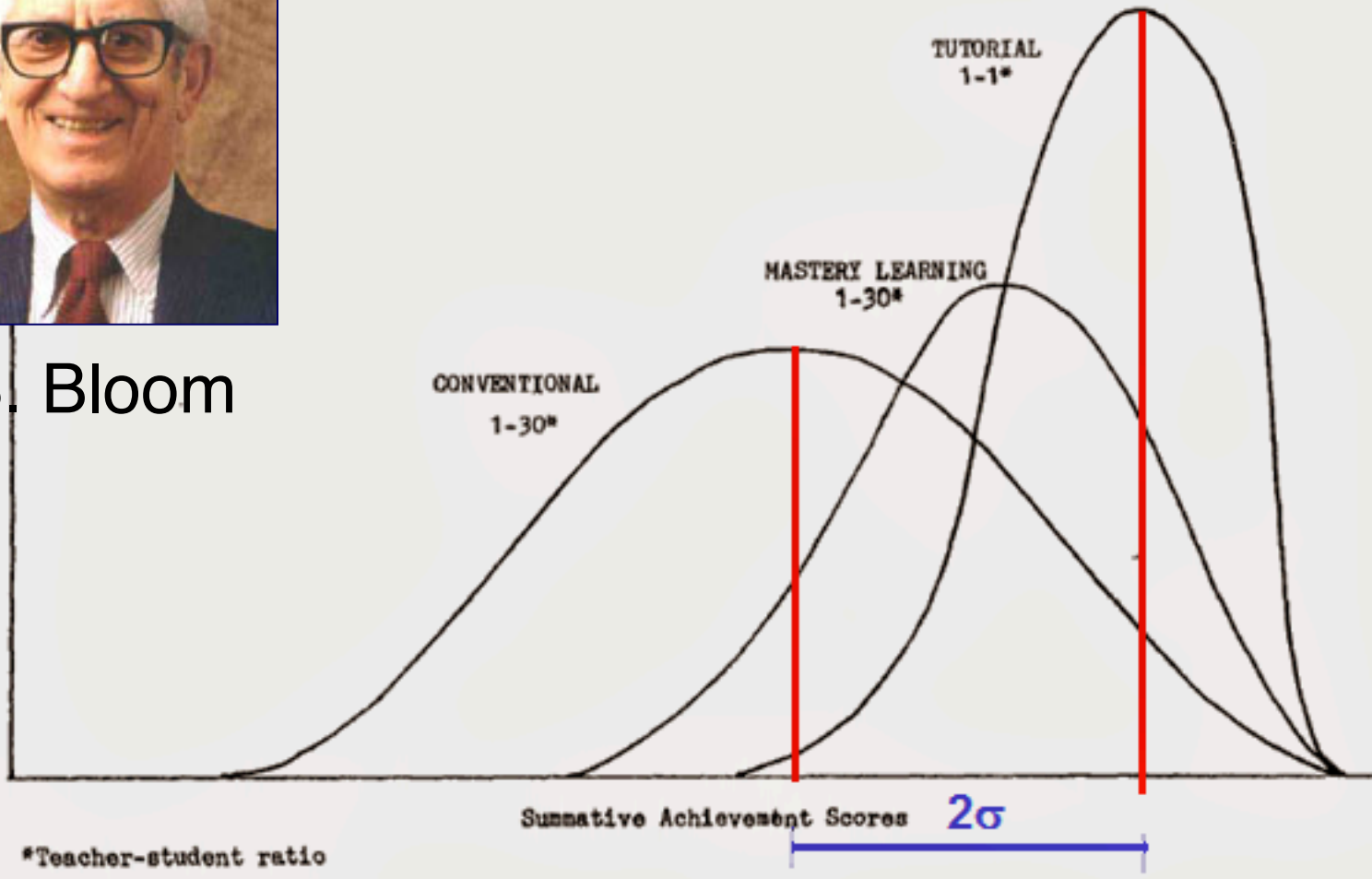
**Intermediate-test:** Did the learner reach the objectives of this module ?

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

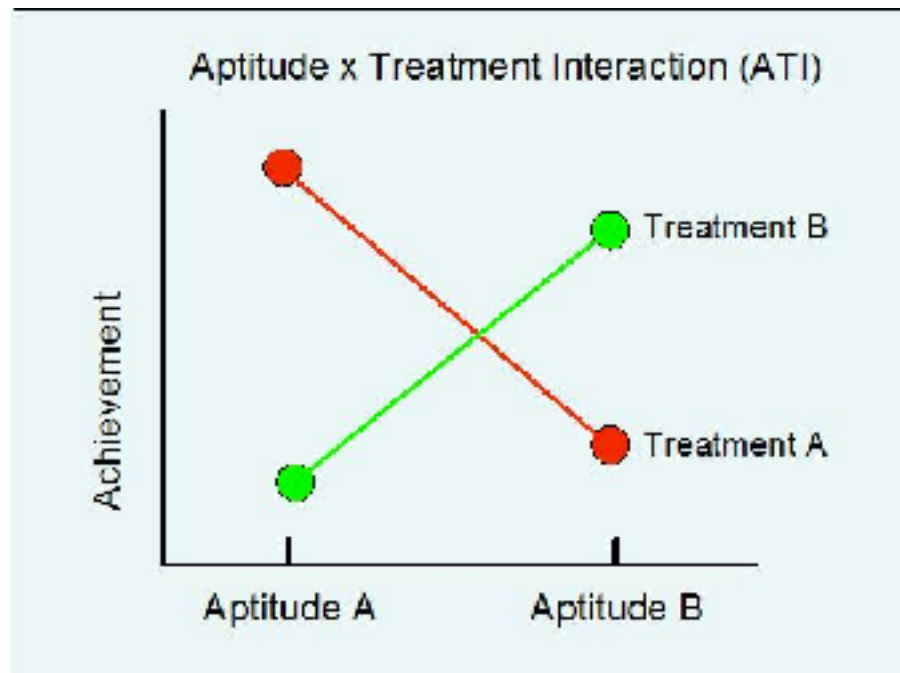
FIGURE 1. Achievement distribution for students under conventional, mastery learning, and tutorial instruction.



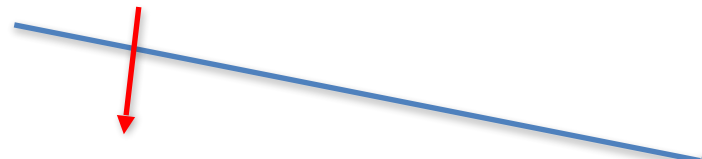
B. Bloom



Bloom, B. (1984). "The 2 Sigma Problem: The Search for Methods of Group Instruction as Effective as One-to-One Tutoring", *Educational Researcher*, 13:6(4-16).



### Aptitude- Treatment Interactions:



The effect of a pedagogical **method** varies for different learners **profiles**

The effect of a pedagogical **method**  
varies for different learners **profiles**

**CHAPTER 9**

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

# Mastery Learning Efficiency

## Lisp Programming Tutor

- Cognitive Mastery vs. Fixed Curriculum:
  - 40% more problems
  - 14% more time
  - 25% greater accuracy on post-test
  - 570% increase in mastery
  - Effect size:  $d=0.65$

Corbett, A. (2001). Cognitive computer tutors: Solving the two-sigma problem. In M. Bauer, P. J. Gmytrasiewicz, & J. Vassileva (Eds.), *Proceedings of the 8<sup>th</sup> International Conference on User Modeling, UM 2001* (pp. 137-147). Springer Berlin Heidelberg. doi:10.1007/3-540-44566-8\_14

# Evaluations of Intelligent Tutoring Systems

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

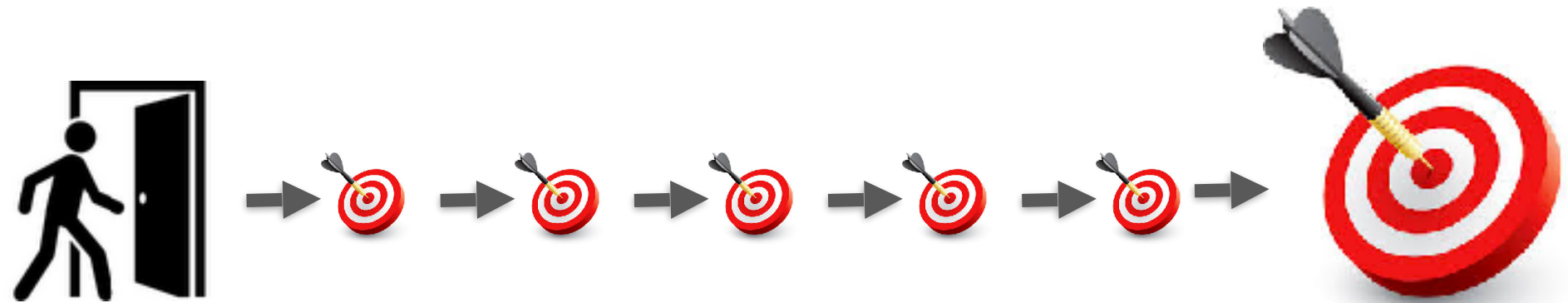
(Pane et al., 2013)

- 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

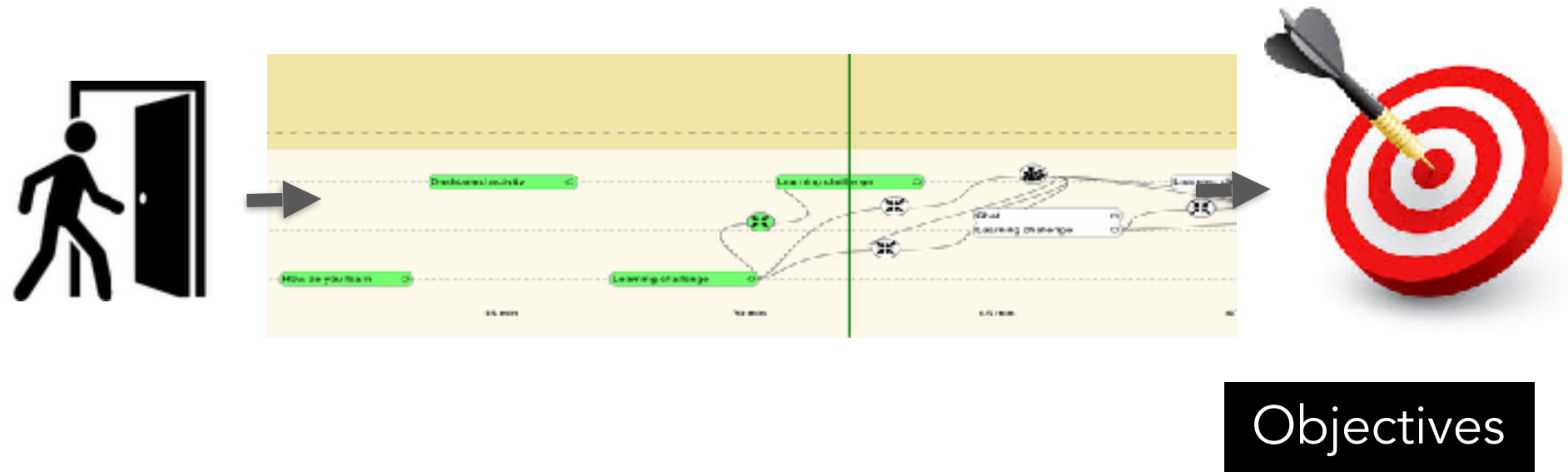
(Kulik & Fletcher, 2015; Ma, Adesope, Nesbit, & Liu, 2014; Steenbergen-Hu & Cooper, 2013; 2014)

# Mastery learning



Objectives

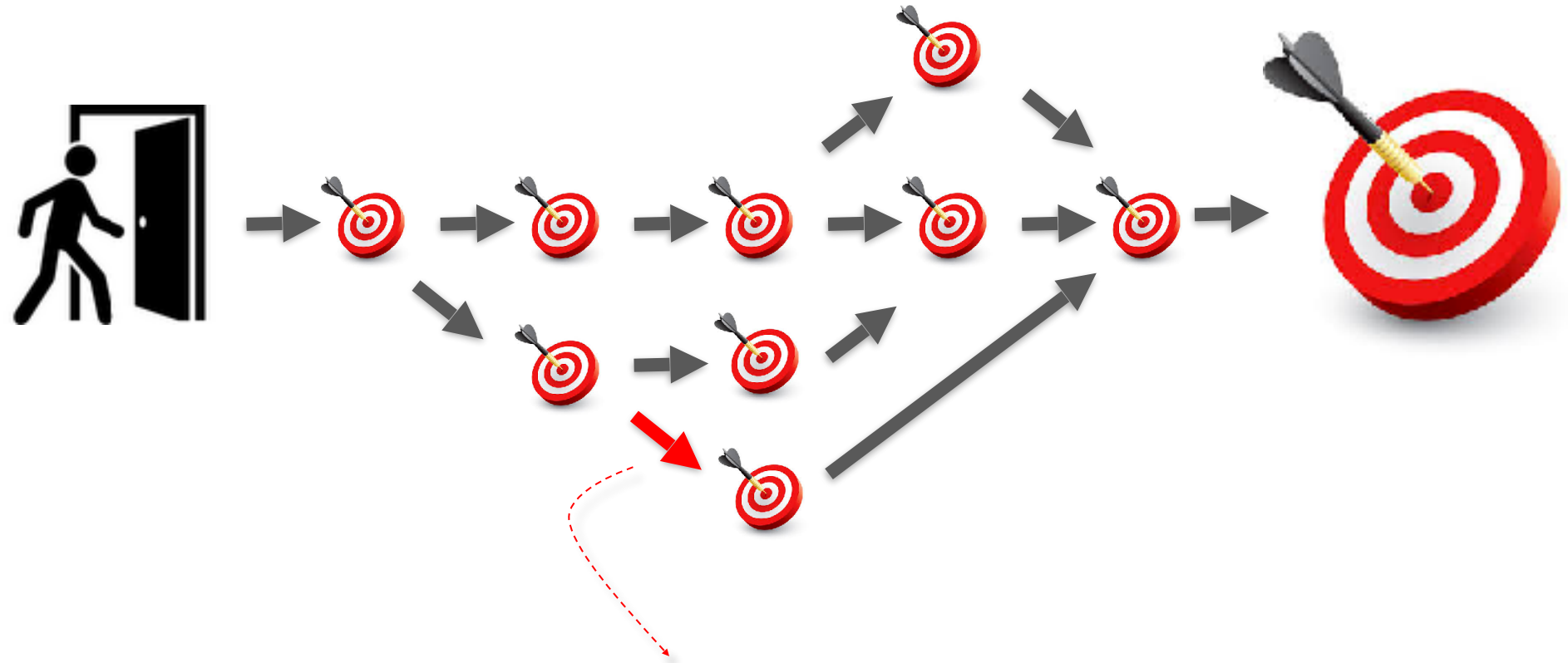
# Mastery learning & Orchestration Graphs



Why is an activity  $a_1$  useful for an activity  $a_2$ ?



# Mastery learning



Why is an activity  $a_1$  useful for an activity  $a_2$ ?

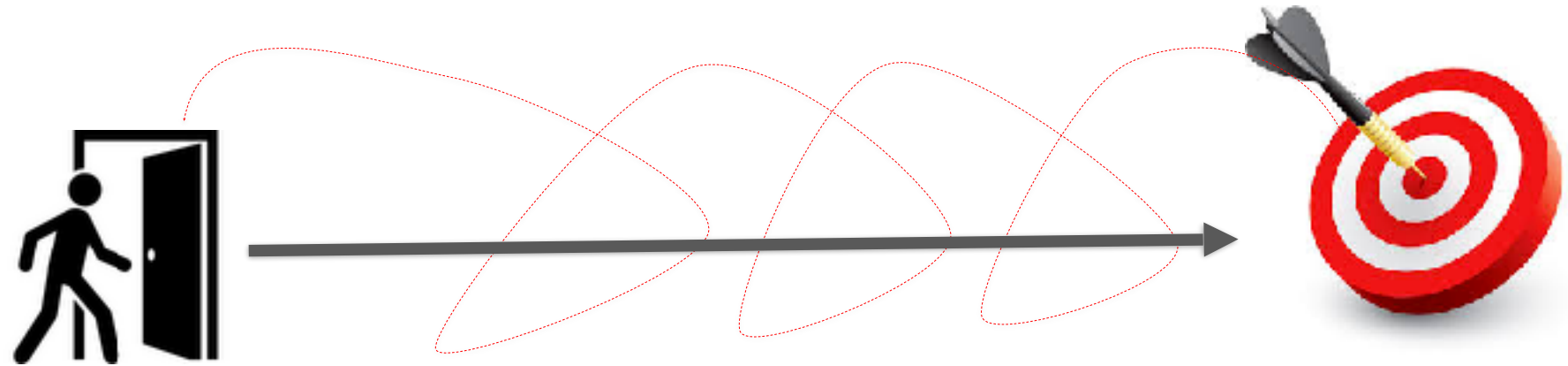
# Why is an activity $a_1$ useful for an activity $a_2$ ?

**Pre-requisite:** {skills- $a_1$ } is a subset of {skills- $a_2$ }

Pre-requisites are common sense: You need to be able to do  $5+7$  ( $a_i$ ), before trying  $25+37$  ( $a_j$ ).

Nonetheless, a high portion of failure is explained the accumulation of small gaps in pre-requisites.

Mastery learning focused especially on this sequencing



What are they able to do  
at the beginning

Pre-Requisites

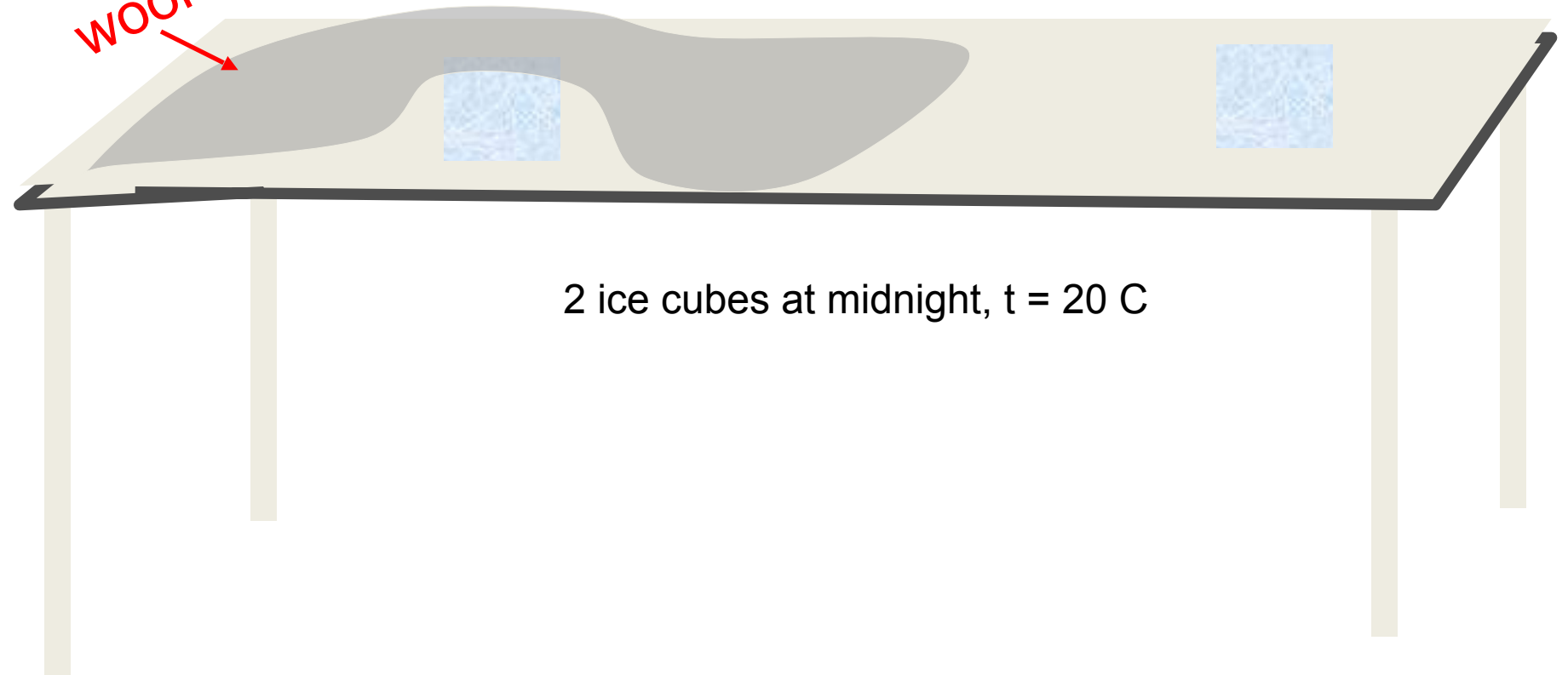
What should learners be able  
to do at the end ?

(which they could not do  
at the beginning)

Objectives

# Prior Knowledge can be wrong

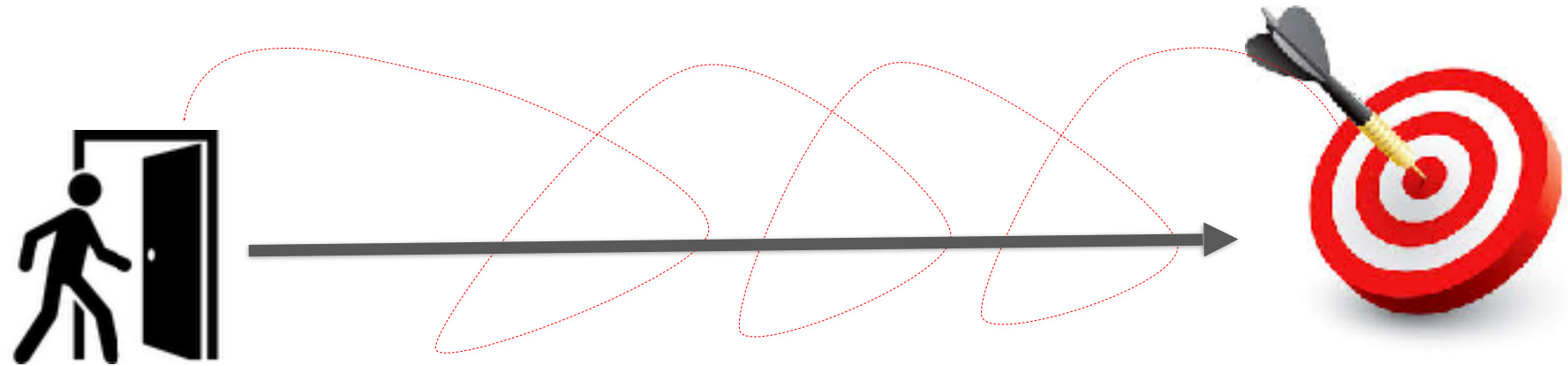
wool



2 ice cubes at midnight,  $t = 20\text{ C}$

Which ice cube will melt faster ?

# Instructional design works backwards :



What are they able to do  
at the beginning

Pre-Requisites

Pre-Representations

What should learners be able  
to do at the end ?

(which they could not do  
at the beginning)

Objectives

Why is an activity  $a_1$  useful for an activity  $a_2$ ?

Advance organizer :

$a_1$  pre-activates structures for  $a_2$

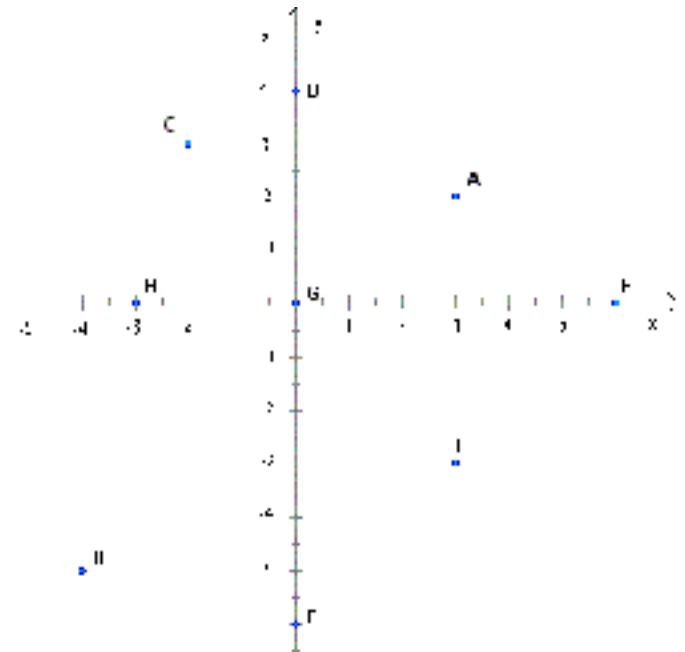
Which fruit do people eat in winter ?

Why is an activity  $a_1$  useful for an activity  $a_2$ ?

Advance organizer :

$a_1$  pre-activates structures for  $a_2$

	1	2	3	4	5	6	7	8	9	10
A										
B		■	■			■		■		
C								■		
D	■							■		
E		■	■	■	■					
F									■	■
G		■	■	■	■					
H										
I		■					■			■
J		■								



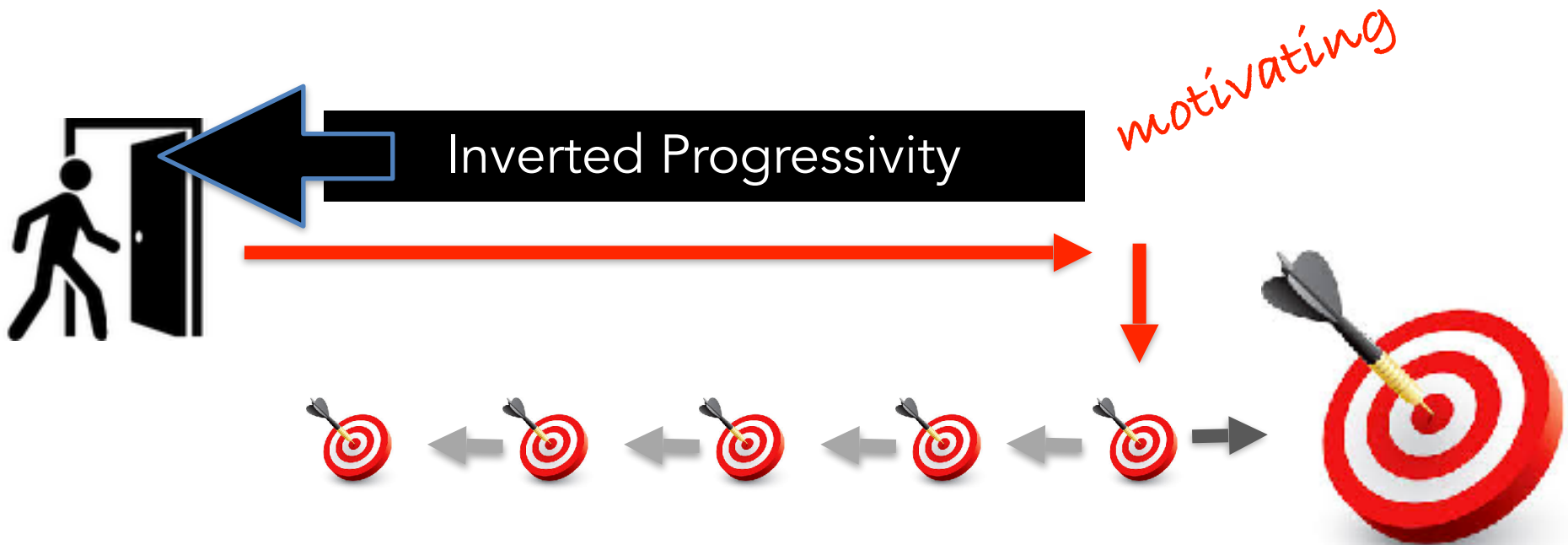
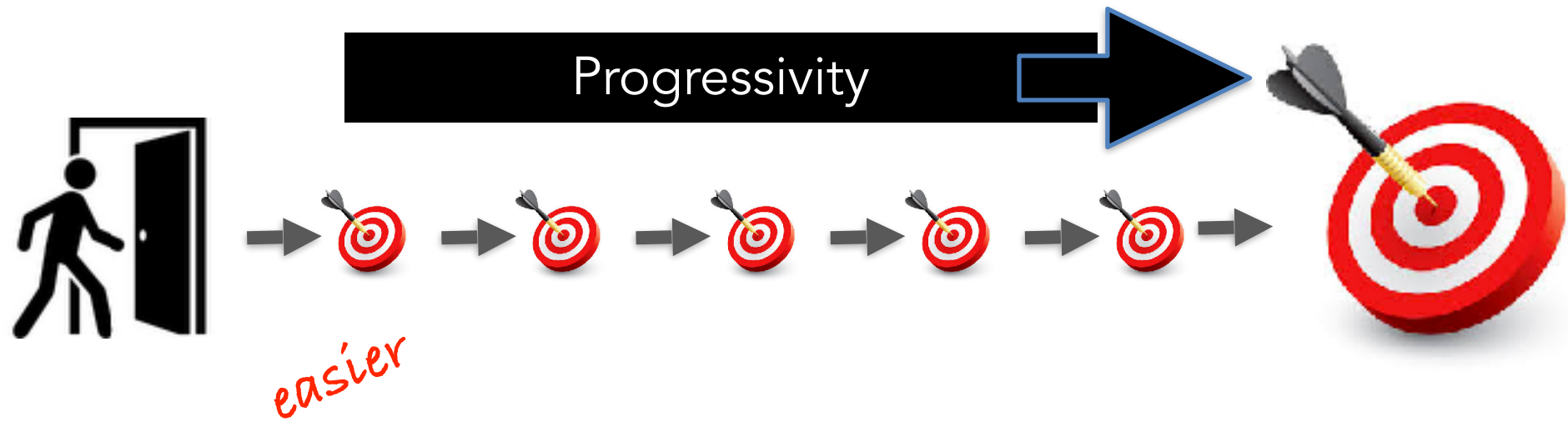




$a_1$

« Let me show you an example of what you will be able to do at the end of this course »





## Types of Motivators



### Extrinsic

- badges
- competition
- fear of failure
- fear of punishment
- gold stars
- money
- points
- rewards
- ...

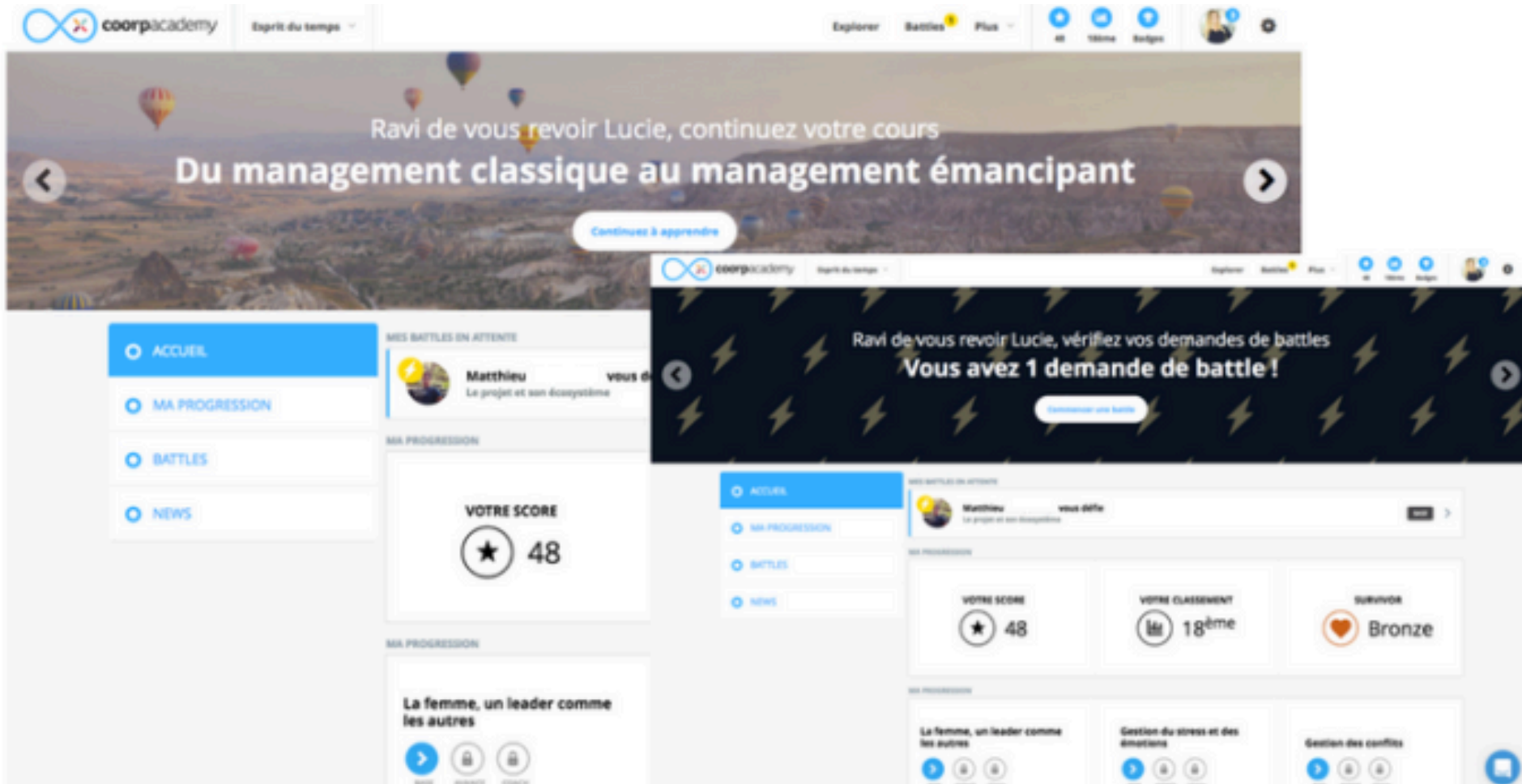


### Intrinsic

- autonomy
- belonging
- curiosity
- love
- learning
- mastery
- meaning
- ...

# Gamification

(even in corporate training)





Gamification

Why is an activity  $a_1$  useful for an activity  $a_2$ ?

Motivation:

$a_1$  frustrate learners so that they want to do  $a_2$

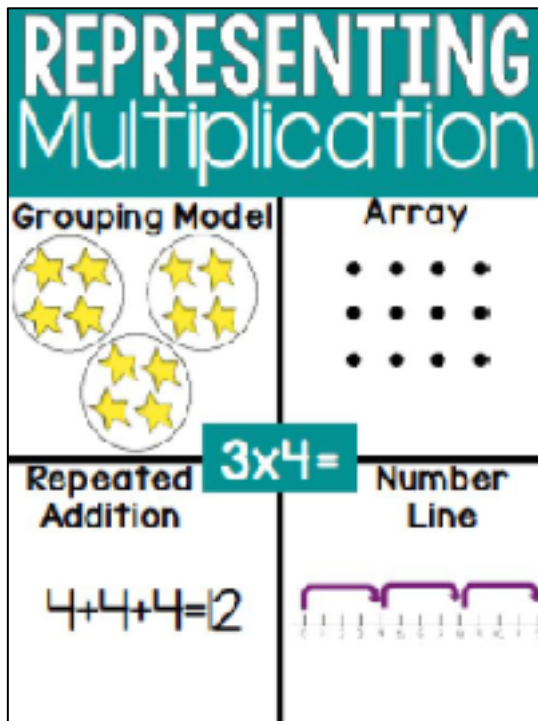
« what you learned before enabled you to solve problems so far, but here are new problems »



Why is an activity  $a_1$  useful for an activity  $a_2$ ?

## Shift Representation:

$a_2$  represents contents differently from  $a_1$



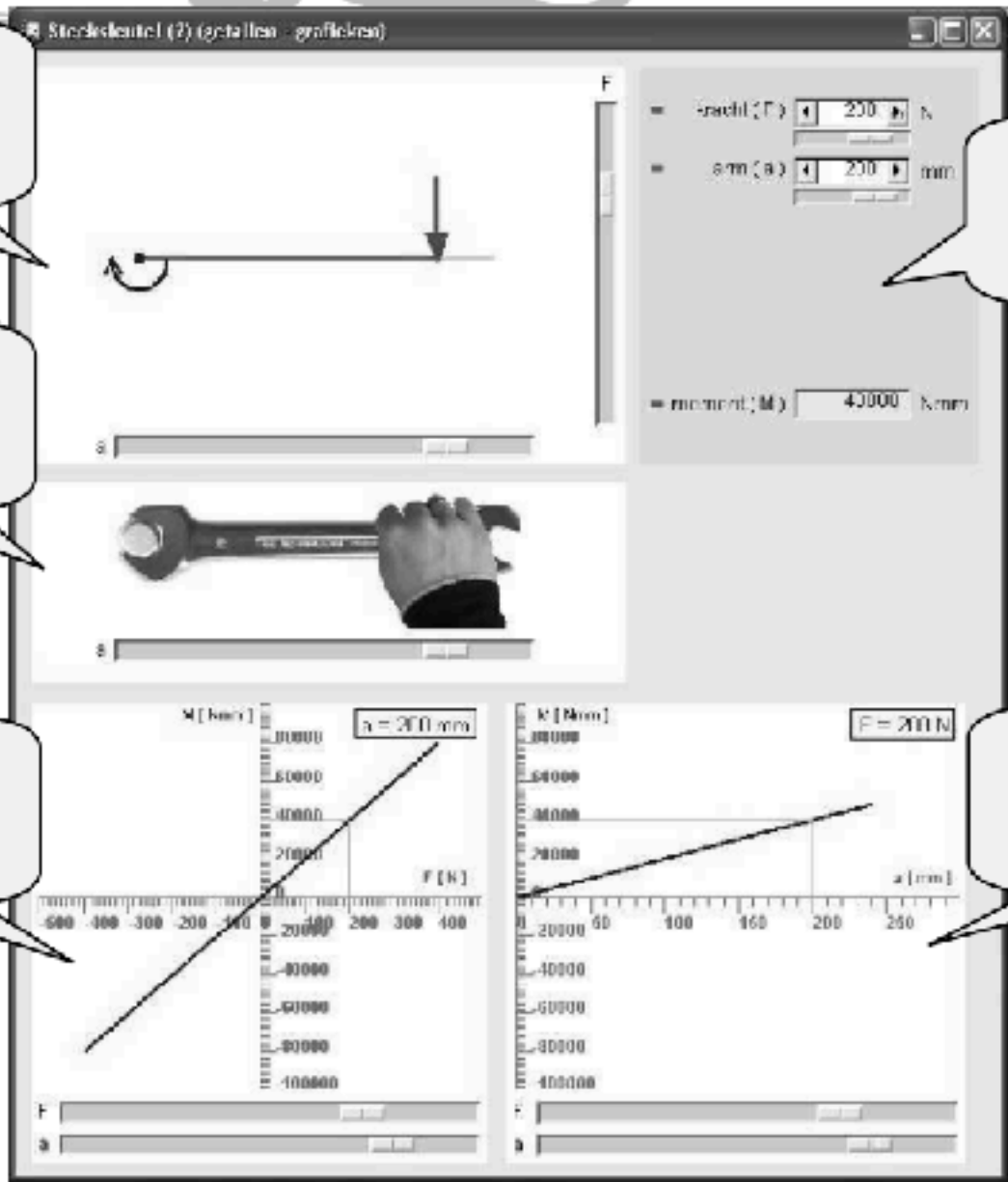
There is a large body of empirical studies that show that is beneficial for learners to **switch between multiple representations.**



Diagrammatic representation

Context representation

Graph



Numerical representation

Graph

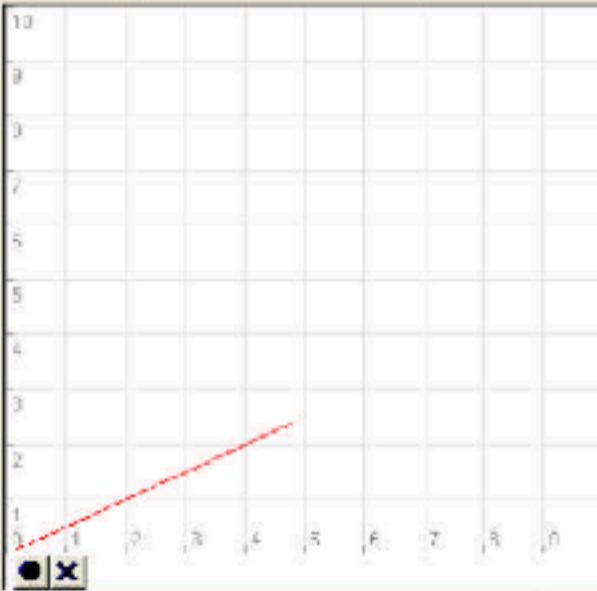
1 Velocity scooter (without controls)

uniform varying motion


scooter starting from stand still

initial velocity  $v(0)$   m/s


acceleration  $a$   m/s<sup>2</sup>



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 0 m/s in 1 seconds.

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

You may try two times.

Start Close

The screenshot displays a software interface with several windows:

- Graph: P and N v Time:** A line graph showing the population of N (green) and P (red) over time. The y-axis is labeled 'Population of individuals' and ranges from 0 to 40. The x-axis is labeled 'Time' and ranges from 0 to 100. The green curve (N) starts at 0, peaks at approximately 35 around time 20, and then fluctuates. The red curve (P) starts at 0, peaks at approximately 10 around time 25, and then fluctuates.
- Animation: Change(X) and X:** A window showing a bar chart with a red bar at the bottom and a green bar at the top. The red bar is labeled 'PACT: 0.076' and the green bar is labeled 'PACT: 0.3754'. The x-axis is labeled 'Field'.
- Table: N and P:** A table showing the population of N and P over time. The y-axis is labeled 'Time' and ranges from 0 to 70. The x-axis is labeled 'N' and 'P'. The table contains the following data:

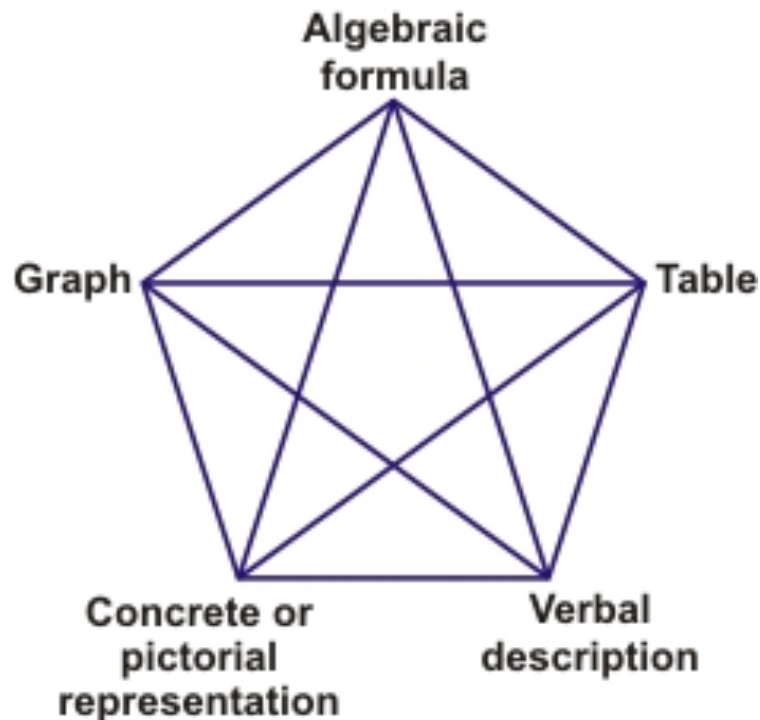
Time	N	P
0	200	5.0
5	284	5.4
10	371	6.9
15	349	8.5
20	266	10.2
25	186	10.9
30	136	10.5
35	113	9.7
40	109	8.7
45	120	7.9
50	144	7.1
55	175	7.1
60	218	7.7
65		
70		

- Simple Equation:** A window showing the simple mathematical model. The equation is  $dN/dt = Potential(N) - Mortal(N)$ . The parameters are:  $dN/dt = 0.38$ ,  $Potential(N) = 5.73$ , and  $Mortal(N) = 5.35$ . The equation is also shown for P:  $dP/dt = Potential(P) - Mortal(P)$  with parameters:  $dP/dt = 0.11$ ,  $Potential(P) = 0.48$ , and  $Mortal(P) = 0.37$ . A note states: "NB If P density was 0, Mortal(P) would be 0 and the behaviour would be the same as single species limited growth."
- Chart: Mort and dP:** A bar chart showing the mortality of N (green) and P (blue) over time. The y-axis is labeled 'Mortality' and ranges from 0 to 1.0. The x-axis is labeled 'Field'.
- Phasorplot: P v N:** A phase plot showing the relationship between P and N. The y-axis is labeled 'P' and ranges from 0 to 15. The x-axis is labeled 'N' and ranges from 10 to 40. The plot shows a closed loop, indicating a limit cycle.
- Animated N and P:** A window showing the population of N and P over time. The x-axis is labeled 'Time' and ranges from 0 to 23. The y-axis is labeled 'Population' and ranges from 0 to 40. The plot shows a series of small icons representing the population of N and P over time.
- Calculator:** A standard Windows calculator window is open in the bottom right corner.

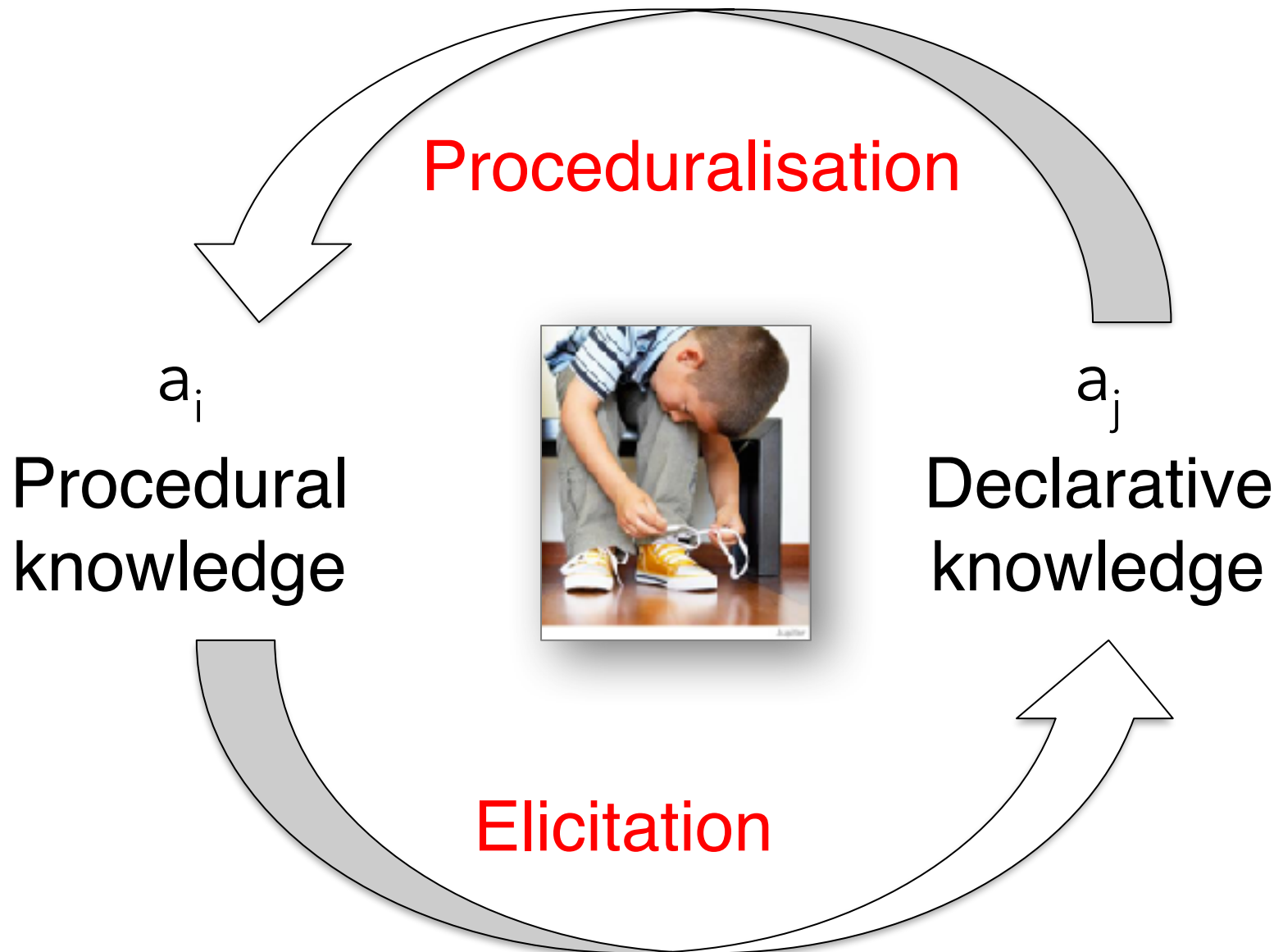
Why is an activity  $a_1$  useful for an activity  $a_2$ ?

## Shift Representation:

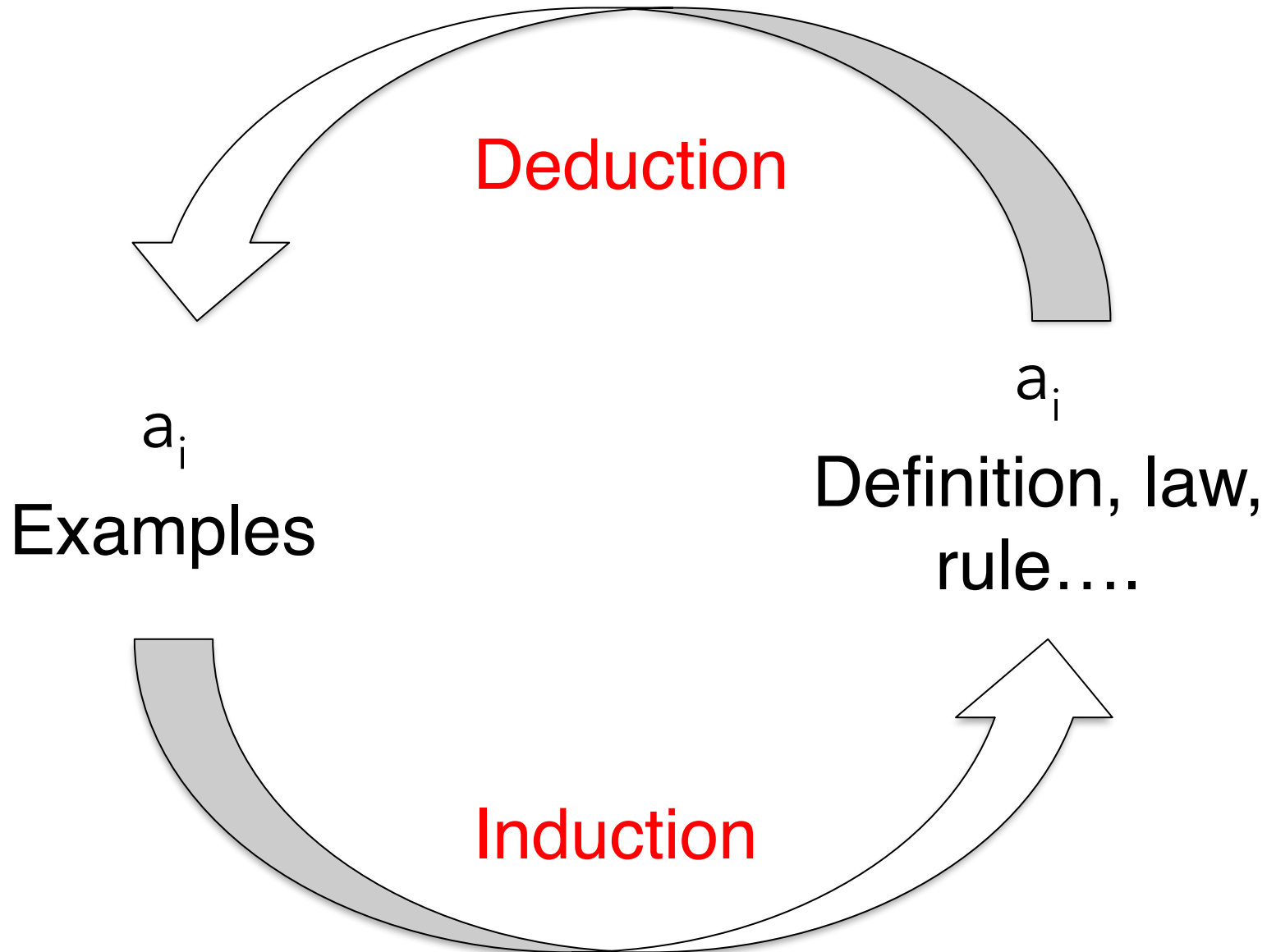
$a_2$  represents contents differently from  $a_1$



Why is an activity  $a_1$  useful for an activity  $a_2$ ?



Why is an activity  $a_1$  useful for an activity  $a_2$ ?



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 27 CHF

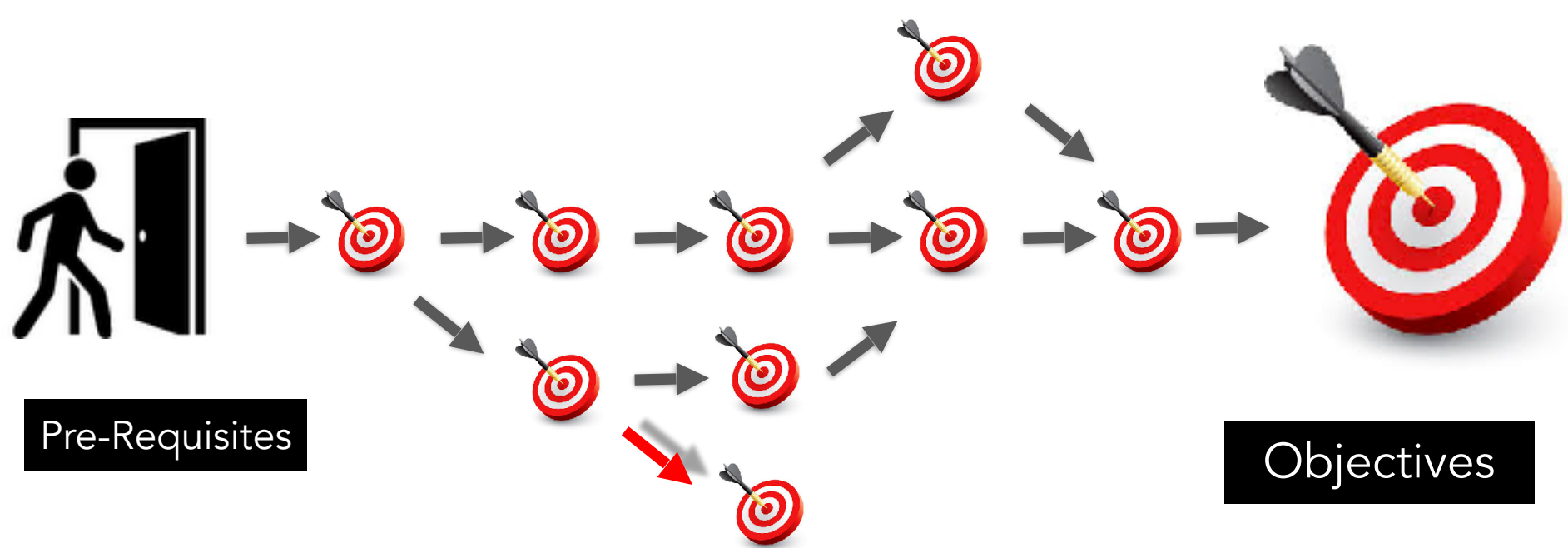
$a_i$

A man walks 1 km in 5 min.

How many km can he walk in 27 min



**Transfer**



- Pre-prerequisite
- Advance Organiser
- Shift representations
- Intrinsic/Extrinsic Representations
- Induction/Deduction
- Proceduralisation/Elicitation
- Transfer



# Examples of **exam** questions

- ① In which ways does this learning technology correspond to a behaviorist approach ?
- ② For which kind of learners is it relevant to choose inverted progressivity ?