Andrew Lampinen

FriSem, 5/11/2018

How do humans learn so quickly and effectively?

Transfer across domains

[Lampinen, Hsu, and Mcclelland, 2017, Hansen, Lampinen, Suri, and McClelland, 2017]

Transfer across domains

[Lampinen et al., 2017, Hansen et al., 2017]

Transfer across domains

category's name:	its objects:	its morphisms:
Set	sets	functions
Group	groups	group homomorphisms
Top	topological spaces	continuous functions
Vect_k	vector spaces over a field, k	linear transformations
Meas	measurable spaces	measurable functions
Poset	partially ordered sets	order-preserving functions
Man	smooth manifolds	smooth maps

[Lampinen et al., 2017, Hansen et al., 2017]

Background

- Not a new idea: transfer has been called a central component of “why we’re so smart” [Gentner, 2003], and an important source of new ideas [Gick and Holyoak, 1980].

Background

- Not a new idea: transfer has been called a central component of "why we're so smart" [Gentner, 2003], and an important source of new ideas [Gick and Holyoak, 1980].
- However, this perspective has been criticized!

Background

- Not a new idea: transfer has been called a central component of “why we’re so smart” [Gentner, 2003], and an important source of new ideas [Gick and Holyoak, 1980].
- However, this perspective has been criticized!
    - “Significant transfer is probably rare and accounts for very little human behavior. ... We generally do what we have learned to do and no more.” [Detterman, 1993]

Background

- Not a new idea: transfer has been called a central component of “why we’re so smart” [Gentner, 2003], and an important source of new ideas [Gick and Holyoak, 1980].
- However, this perspective has been criticized!
    - “Significant transfer is probably rare and accounts for very little human behavior. ... We generally do what we have learned to do and no more.” [Detterman, 1993]
- How can we reconcile these different viewpoints?

Transfer speed

I think there's a neglected variable: speed of transfer.

[Lampinen et al., 2017]

Transfer speed

I think there's a neglected variable: speed of transfer.

Fast	Slow
• one (or a few) examples explicitly shown	• learning gradually through many interactions

[Lampinen et al., 2017]

Transfer speed

I think there's a neglected variable: speed of transfer.

Fast	Slow
• one (or a few) examples explicitly shown • transfer requires explicit awareness of analogy	• learning gradually through many interactions • transfer may or may not be explicit

[Lampinen et al., 2017]

Transfer speed and outcomes

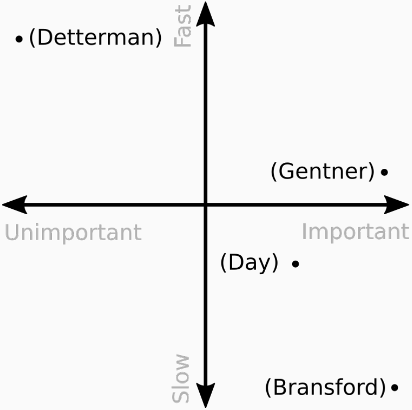
For example, imagine you are about to teach someone to play guitar. Who do you think would likely perform better:

- a person who has just taken their first piano lesson
- a person who has been playing piano for years

Transfer speed and opinions

Transfer speed helps reconcile why some people think transfer is important, while others don't.



Slow transfer between tasks is important.

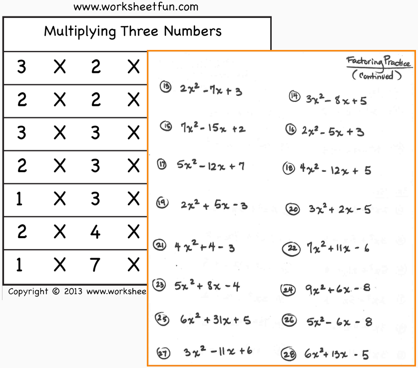
Is abstraction transfer?

We often progress from procedural to more explicit, abstract, or formal knowledge:

www.worksheetfun.com

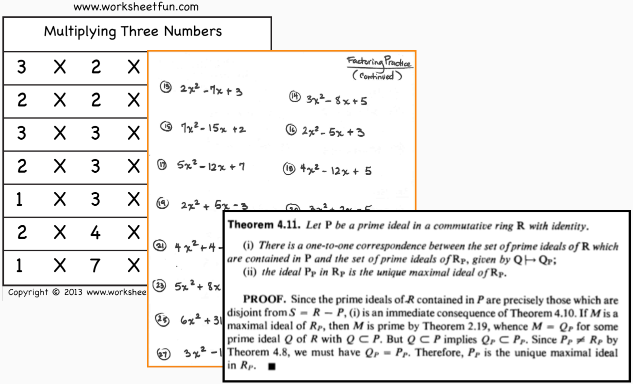
Multiplying Three Numbers					
3	X	2	X	2	=
2	X	2	X	8	=
3	X	3	X	5	=
2	X	3	X	7	=
1	X	3	X	9	=
2	X	4	X	8	=
1	X	7	X	4	=

We often progress from procedural to more explicit, abstract, or formal knowledge:



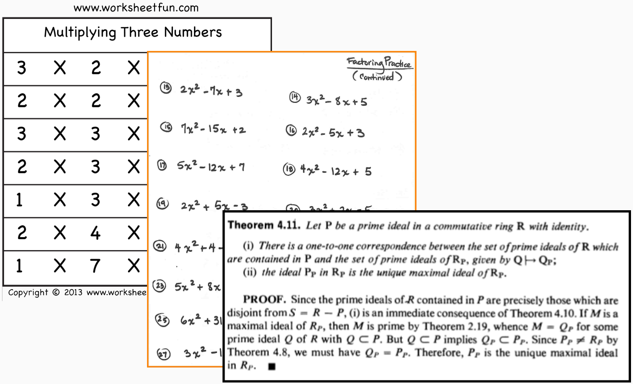
Is abstraction transfer?

We often progress from procedural to more explicit, abstract, or formal knowledge:



Is abstraction transfer?

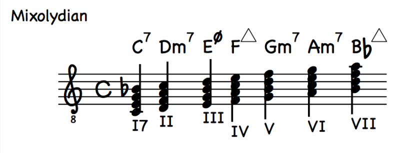
We often progress from procedural to more explicit, abstract, or formal knowledge:



These are related tasks that share some structure! Is it useful to think about this from the perspective of transfer?

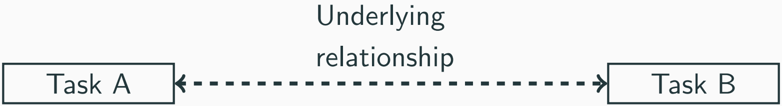
Is abstraction transfer?

Yet there are also times when procedural and formal knowledge seem partially dissociable:

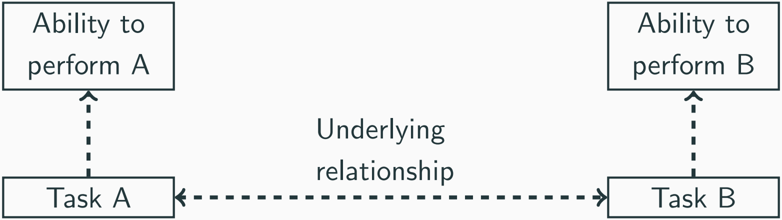


When and how does procedural knowledge transfer to more explicit, abstract, or formal knowledge?

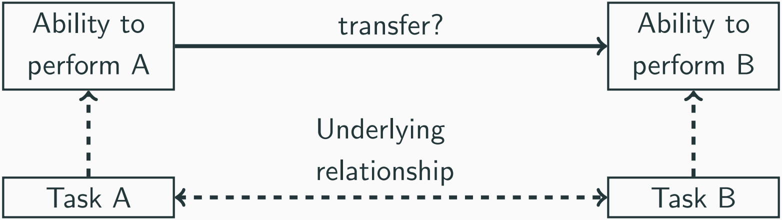
A diagram of potential phenomena of interest



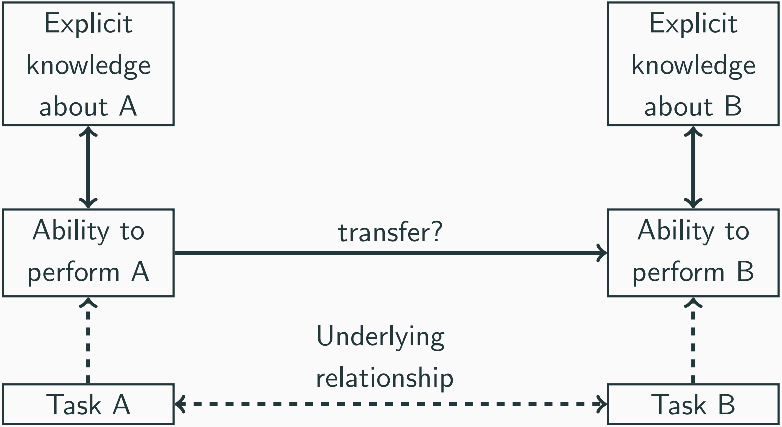
A diagram of potential phenomena of interest



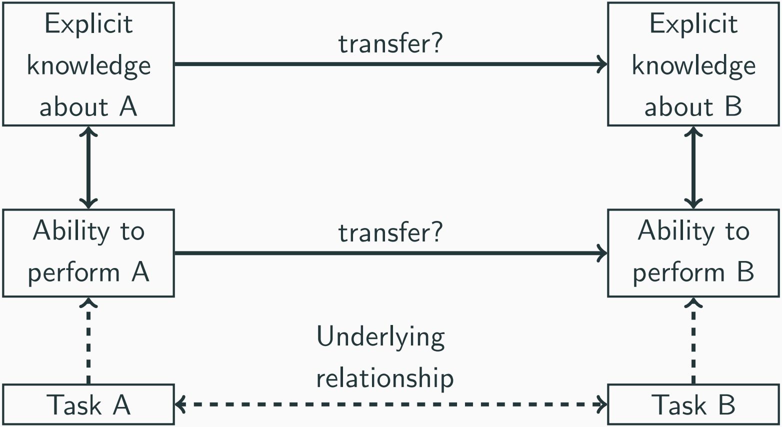
A diagram of potential phenomena of interest



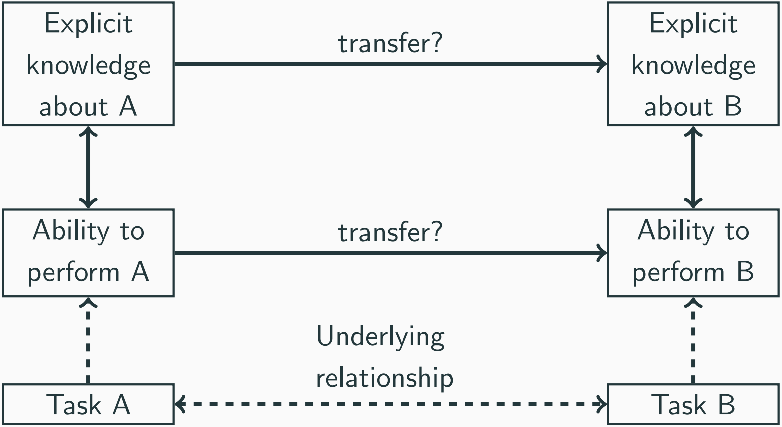
A diagram of potential phenomena of interest



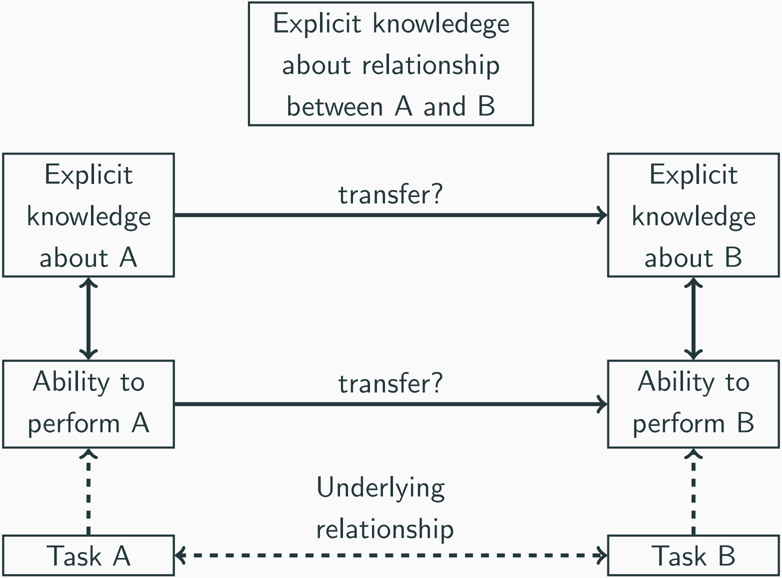
A diagram of potential phenomena of interest



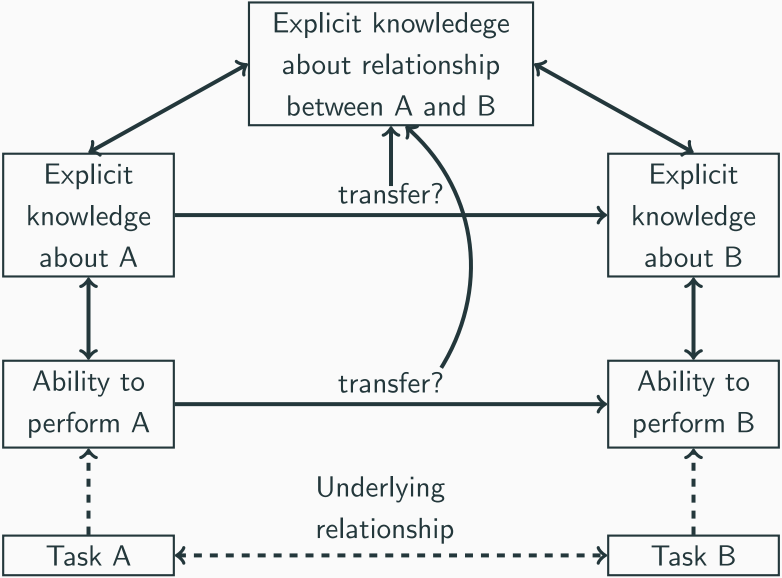
A diagram of potential phenomena of interest



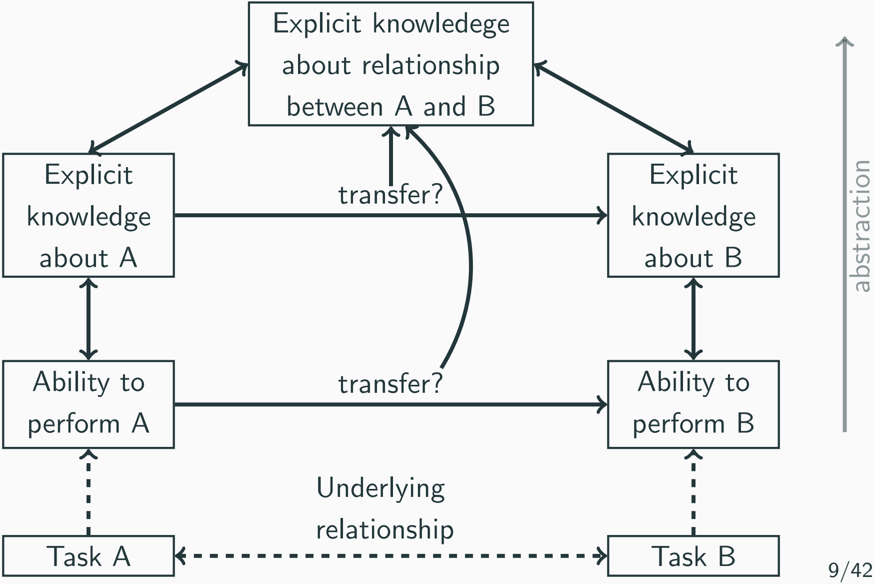
A diagram of potential phenomena of interest



A diagram of potential phenomena of interest

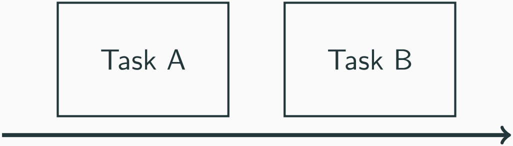


A diagram of potential phenomena of interest

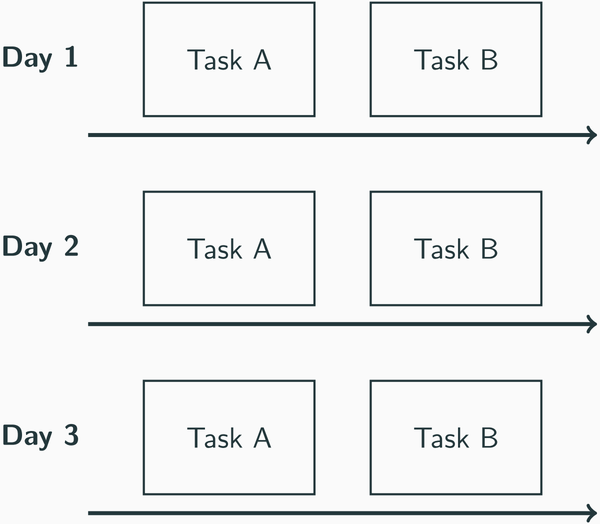


Experiment design

Experiment design



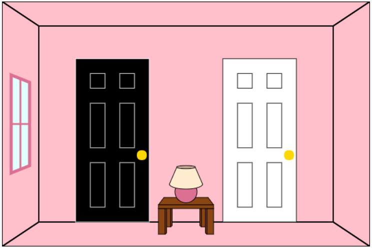
Experiment design

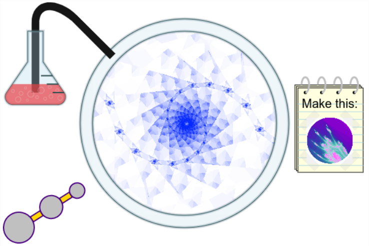


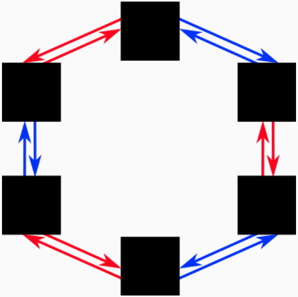
Tasks

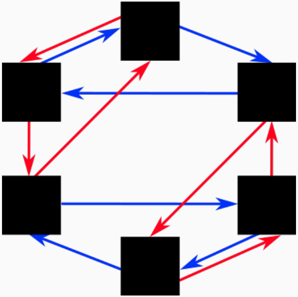
https://web.stanford.edu/~lampinen/mturk/il/web/experiment_mockup.html

Tasks

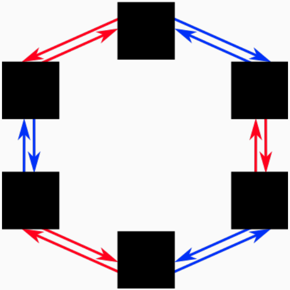
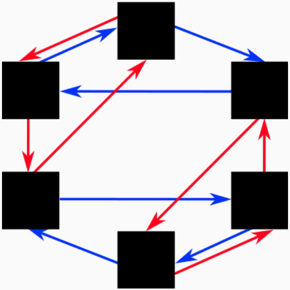




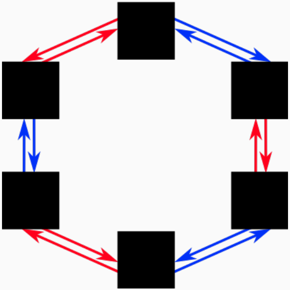
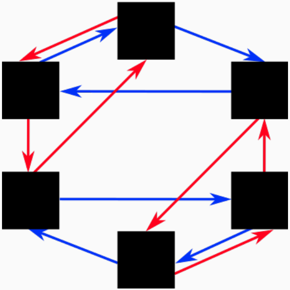




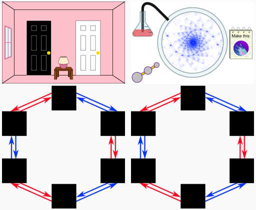
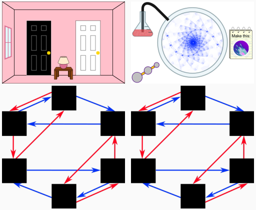
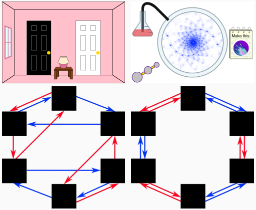
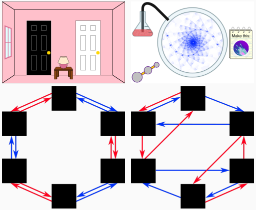
Structures

Flipping	Cycling
	

Structures

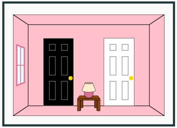
Flipping	Cycling
	

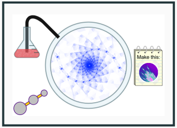
- Balanced for average path length, number of actions entering/exiting a room, etc.

	Fractal structure	
	Flipping	Cycling
Isomorphic		
Non-isomorphic		

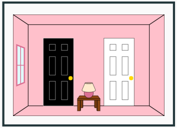
Experiment design

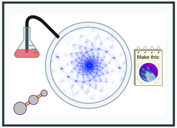
Day 1



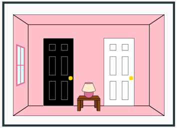


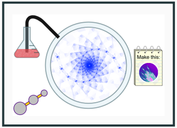
Day 2





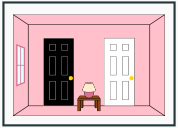
Day 3

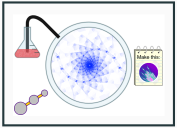




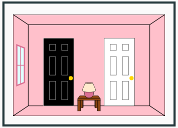
Experiment design

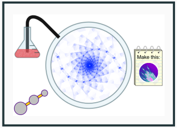
Day 1



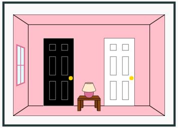


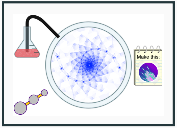
Day 2





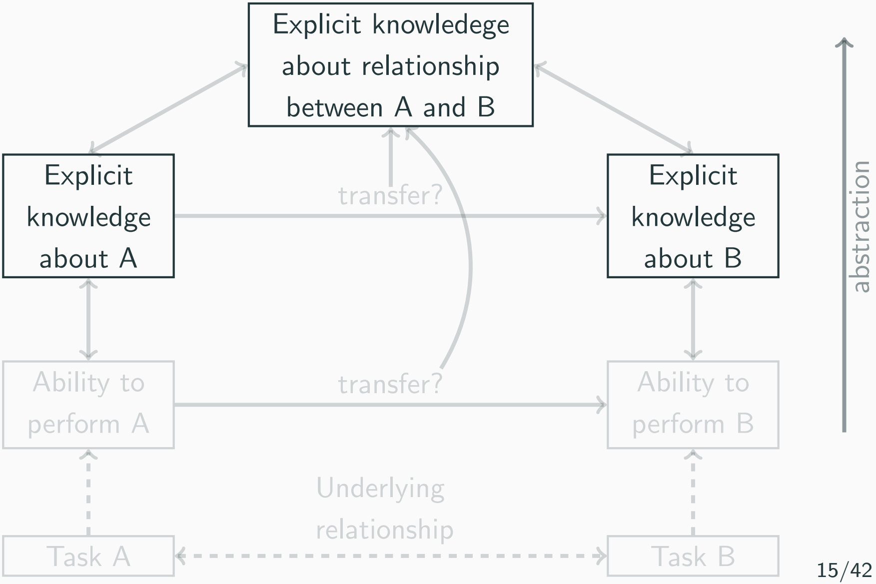
Day 3





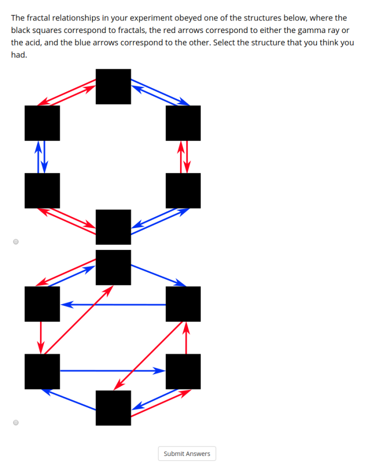
Explicit knowledge?

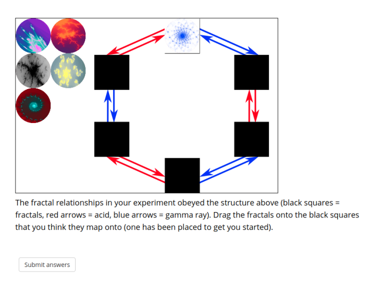
What explicit knowledge?



Explicit knowledge about a single task

Within fractals:



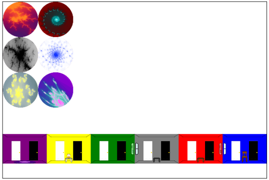


Explicit knowledge about the relationship between tasks

Between tasks:

If the experiments were teaching you the same structure, one of the doors in the spatial navigation experiment would have corresponded to the gamma ray in the fractal experiment, and the other door would have corresponded to the acid. What would be your best guess about the correspondence between the doors and the gamma rays?

- ○ left (black) door = acid, right (white) door = ray
- ◉ left (black) door = ray, right (white) door = acid

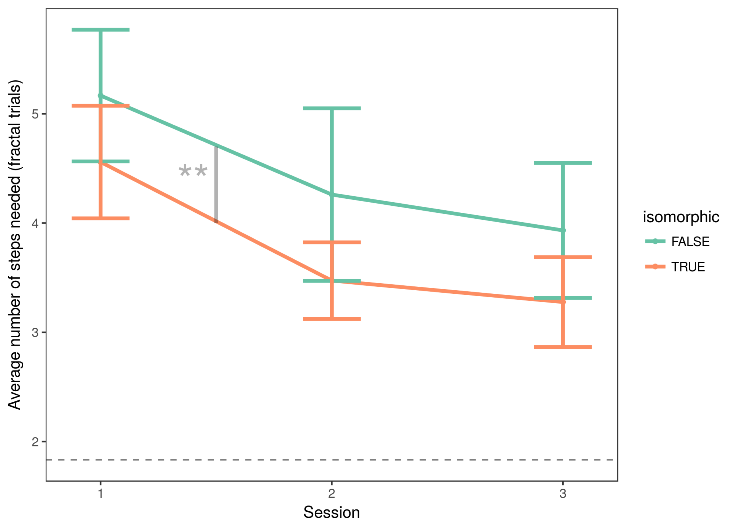


Assuming the experiments were teaching you the same structure, each room in the navigation experiment corresponded to a fractal in the visual patterns experiment. Drag the fractals onto the rooms you think they corresponded to.

Submit answers

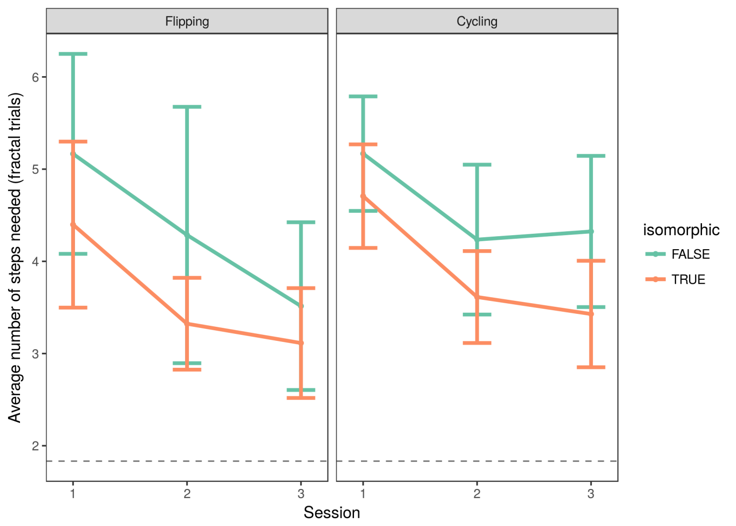
Results

Subjects in isomorphic condition perform better at fractals



Error bars represent 95%-confidence intervals

Subjects in isomorphic condition perform better at fractals

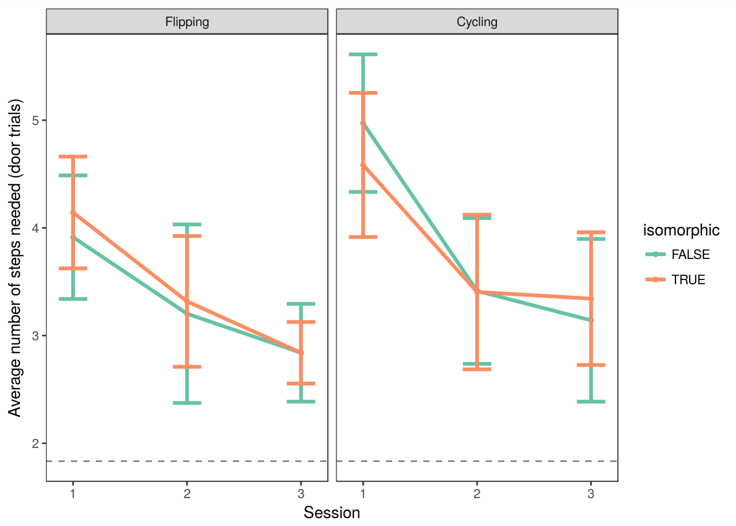


Error bars represent 95%-confidence intervals

Subjects in isomorphic condition perform better at fractals

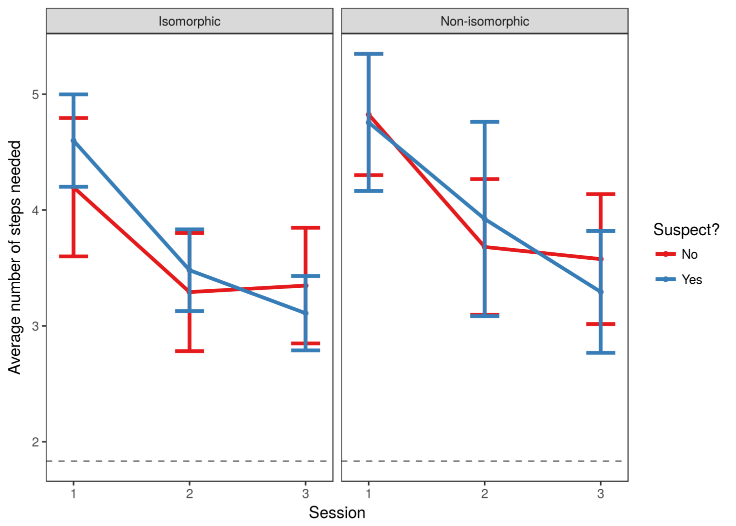
```
Linear mixed model fit by REML
t-tests use  Satterthwaite approximations to degrees of freedom ['lmerMod']
Formula: num_steps_needed ~ isomorphic + session_c + I(session_c^2) +      structure + trial_index_by_type_z +
   Data: num_steps_per_trial_data %>% filter(trial_type == "fractal")
                                                                          I(trial_index_by_type_z^2) +  avg_rt_on_trial + (1 + session_c | subject_id)

REML criterion at convergence: 46673.1

Scaled residuals:
    Min      1Q  Median      3Q     Max
-2.4192 -0.5138 -0.1547  0.2466 14.9688

Random effects:
 Groups     Name        Variance Std.Dev. Corr
 subject_id (Intercept)  1.3336  1.1548
            session_c    0.2308  0.4805   0.00
 Residual               15.2730  3.9081
Number of obs: 8350, groups:  subject_id, 56

Fixed effects:
                             Estimate Std. Error         df t value Pr(>|t|)
(Intercept)                   4.99586    0.29402   67.00000  16.992  < 2e-16 ***
isomorphicTRUE               -0.87140    0.32063   53.00000  -2.718  0.00886 **
session_c                    -0.55866    0.08284   55.00000  -6.744 9.96e-09 ***
I(session_c^2)                0.21058    0.09152 8250.00000   2.301  0.02142 *
structureCycling              0.36442    0.32052   53.00000   1.137  0.26067
trial_index_by_type_z        -0.35697    0.04444 8259.00000  -8.033 1.11e-15 ***
I(trial_index_by_type_z^2)    0.27053    0.04964 8238.00000   5.450 5.18e-08 ***
avg_rt_on_trial              -0.56729    0.02996 5704.00000 -18.937  < 2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

No effect on door performance



Error bars represent 95%-confidence intervals

This may not be conscious



Error bars represent 95%-confidence intervals

This may not be conscious

```
Linear mixed model fit by REML
t-tests use  Satterthwaite approximations to degrees of freedom ['lmerMod']
Formula: num_steps_needed ~ isomorphic * (correspondence_suspect) + session_c +
   Data: es_joined_ns_data %>% filter(trial_type == "fractal")
                                                               I(session_c^2) + structure + trial_index_by_type_z + I(trial_index_by_type_z^2) +

REML criterion at convergence: 46675
                                                                avg_rt_on_trial + (1 + session_c | subject_id)

Scaled residuals:
    Min      1Q  Median      3Q     Max
-2.4227 -0.5136 -0.1547  0.2470 14.9688

Random effects:
 Groups     Name        Variance Std.Dev. Corr
 subject_id (Intercept)  1.3898  1.1789
            session_c    0.2308  0.4805   0.00
 Residual               15.2730  3.9081
Number of obs: 8350, groups:  subject_id, 56

Fixed effects:
                                          Estimate Std. Error         df t value Pr(>|t|)
(Intercept)                                4.99688    0.30253   63.00000  16.517  < 2e-16 ***
isomorphicTRUE                            -0.87178    0.34104   51.00000  -2.556   0.0136 *
correspondence_suspectYes                  0.01273    0.23669   51.00000   0.054   0.9573
session_c                                 -0.55867    0.08284   55.00000  -6.744 9.95e-09 ***
I(session_c^2)                             0.21056    0.09152 8250.00000   2.301   0.0214 *
structureCycling                           0.36663    0.34714   51.00000   1.056   0.2959
trial_index_by_type_z                     -0.35696    0.04444 8259.00000  -8.032 1.11e-15 ***
I(trial_index_by_type_z^2)                 0.27053    0.04964 8238.00000   5.450 5.19e-08 ***
avg_rt_on_trial                           -0.56724    0.02999 5753.00000 -18.914  < 2e-16 ***
isomorphicTRUE:correspondence_suspectYes  -0.01833    0.33967   51.00000  -0.054   0.9572
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

This may not be conscious

- Consciousness is tricky to assess [Newell and Shanks, 2014].

This may not be conscious

- Consciousness is tricky to assess [Newell and Shanks, 2014].
    - Can be altered by the exact wording of question or type of probe.

This may not be conscious

- Consciousness is tricky to assess [Newell and Shanks, 2014].
  - Can be altered by the exact wording of question or type of probe.
- For what it's worth, regressors other than suspecting correspondence give similar results, e.g.:

This may not be conscious

- Consciousness is tricky to assess [Newell and Shanks, 2014].
    - Can be altered by the exact wording of question or type of probe.
- For what it's worth, regressors other than suspecting correspondence give similar results, e.g.:
    - Likert ratings of perception that one task was helpful for the other.

This may not be conscious

- Consciousness is tricky to assess [Newell and Shanks, 2014].
    - Can be altered by the exact wording of question or type of probe.
- For what it's worth, regressors other than suspecting correspondence give similar results, e.g.:
    - Likert ratings of perception that one task was helpful for the other.
    - Likert ratings of similarity between the experiments.

This may not be conscious

- Consciousness is tricky to assess [Newell and Shanks, 2014].
    - Can be altered by the exact wording of question or type of probe.
- For what it's worth, regressors other than suspecting correspondence give similar results, e.g.:
    - Likert ratings of perception that one task was helpful for the other.
    - Likert ratings of similarity between the experiments.
    - Binary split of reports of when they noticed any relationship between the experiments (ever noticed/never noticed).

Evidence of transfer to second task.

Evidence of transfer to second task.

May not require consciousness.

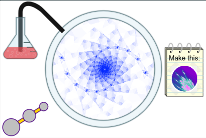
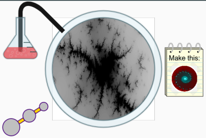
Results: Policies

Policies

- A number of my remaining analyses rely on the idea of a *policy* from reinforcement learning.

Policies

- A number of my remaining analyses rely on the idea of a *policy* from reinforcement learning.
- A policy is a function that maps states to action probabilities:

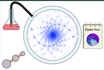
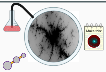
	$P(\text{Acid})$	$P(\text{Ray})$
	0.75	0.25
	0.5	0.5

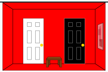
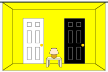
Policy alignment

- How similar is participants behavior between the tasks?

Policy alignment

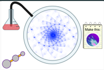
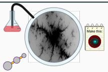
- How similar is participants behavior between the tasks?
- We'll quantify this by looking at similarity between their policies.

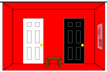
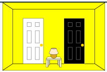
	$P(\text{Acid})$	$P(\text{Ray})$
	0.75	0.25
	0.5	0.5
⋮	⋮	⋮

	$P(\text{Right})$	$P(\text{Left})$
	0.5	0.5
	0.75	0.25
⋮	⋮	⋮

Policy alignment

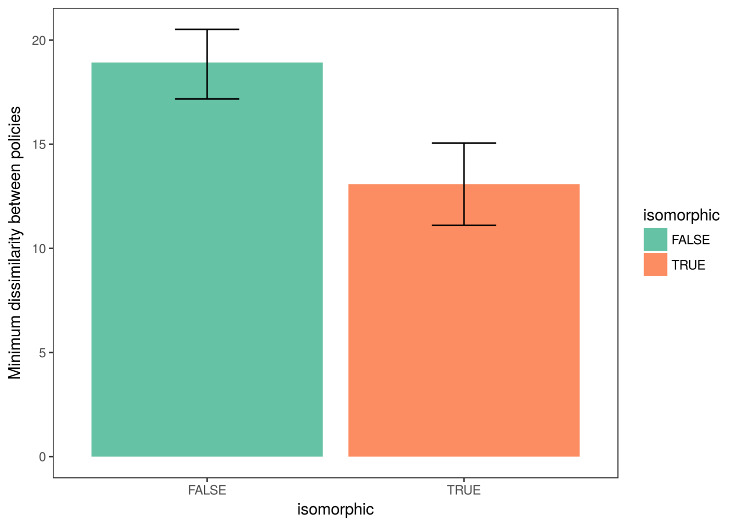
- How similar is participants behavior between the tasks?
- We'll quantify this by looking at similarity between their policies.

	$P(\text{Acid})$	$P(\text{Ray})$
	0.75	0.25
	0.5	0.5
$\vdots$	$\vdots$	$\vdots$

	$P(\text{Right})$	$P(\text{Left})$
	0.5	0.5
	0.75	0.25
$\vdots$	$\vdots$	$\vdots$

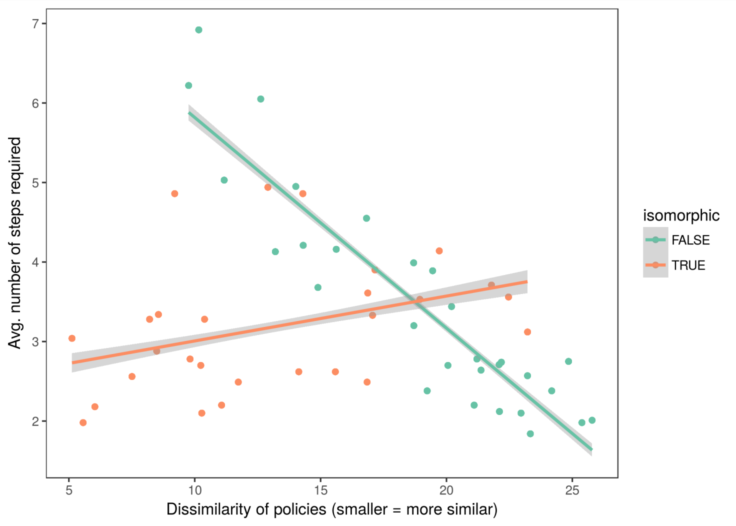
- We'll consider all the structure-preserving mappings between the doors and fractal policies, and see how well the participants' policies align across tasks.

Policy alignment is greater in isomorphic condition



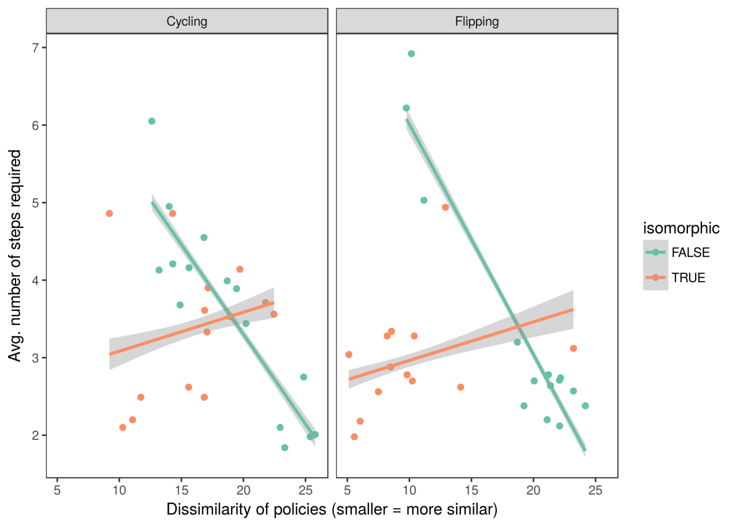
Error bars represent bootstrap 95%-confidence intervals

Policy alignment is related to performance



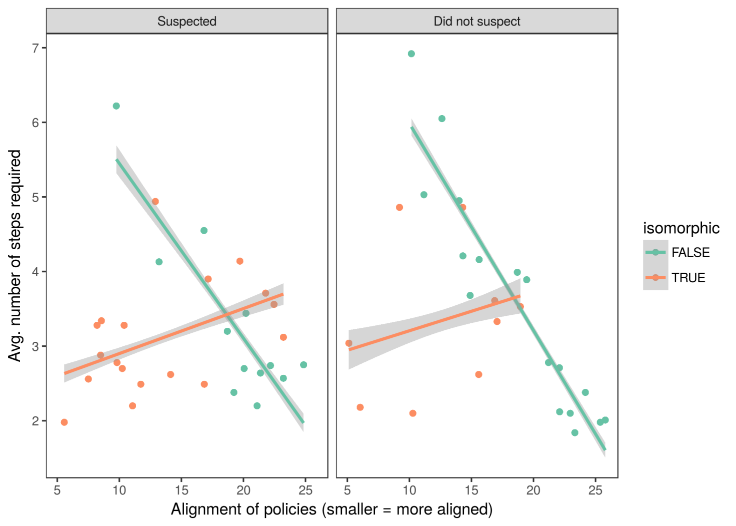
Error bars represent 95%-confidence intervals

Policy alignment is related to performance



Error bars represent 95%-confidence intervals

Not a consciousness effect?



Error bars represent 95%-confidence intervals

Policy alignment strongly affects performance.

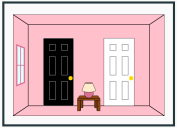
Policy alignment strongly affects performance.

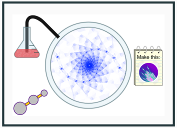
Some subjects may be “trying” to transfer in the non-isomorphic condition, and actually be harmed by it!

Results: Explicit knowledge

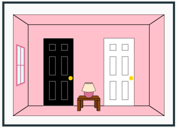
Experiment design

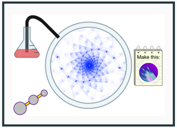
Day 1



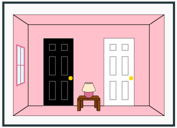


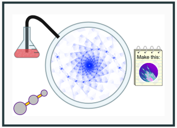
Day 2





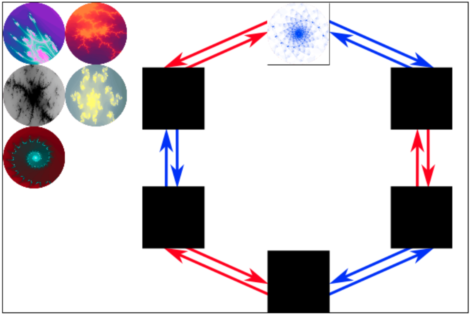
Day 3





Note: we have much less power to detect these effects, because they're generally based off of one answer per question per subject in only the final session.

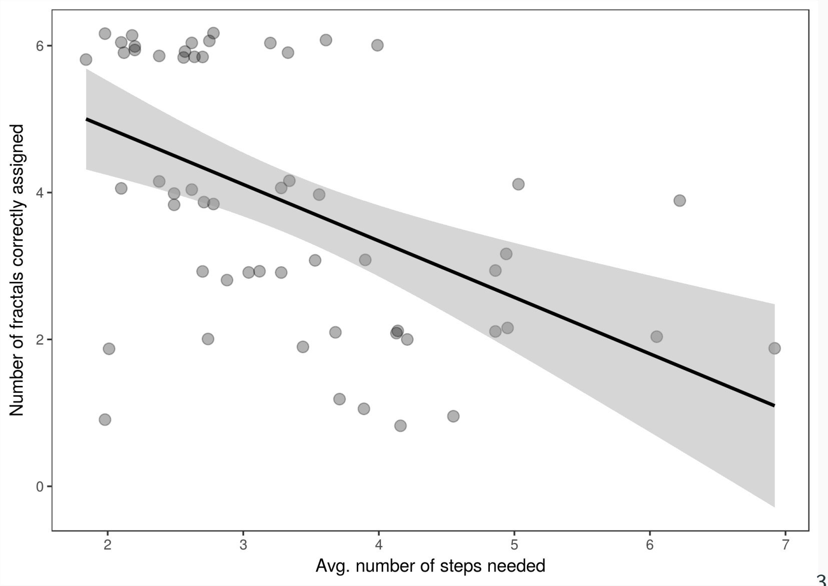
Better performance on task predicts better explicit answers



The fractal relationships in your experiment obeyed the structure above (black squares = fractals, red arrows = acid, blue arrows = gamma ray). Drag the fractals onto the black squares that you think they map onto (one has been placed to get you started).

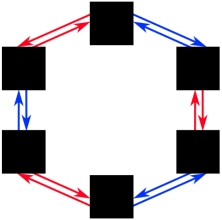
Submit answers

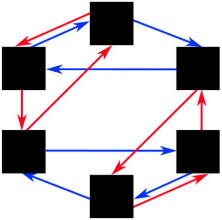
Better performance on task predicts better explicit answers



Isomorphism may influence within-task abstraction

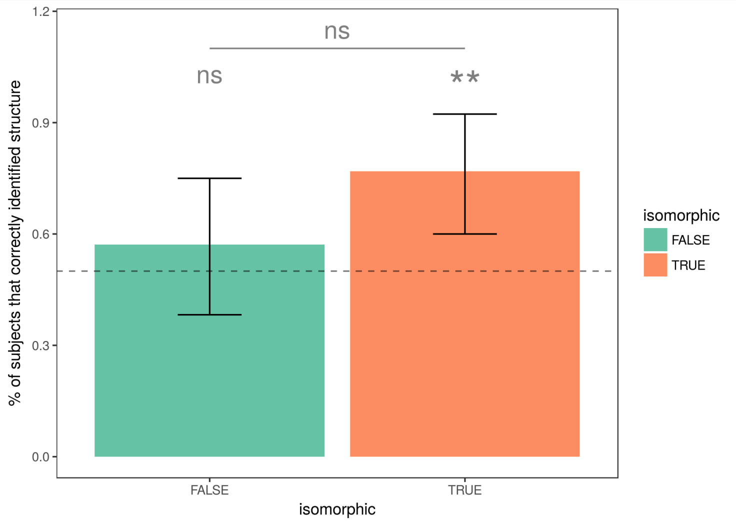
The fractal relationships in your experiment obeyed one of the structures below, where the black squares correspond to fractals, the red arrows correspond to either the gamma ray or the acid, and the blue arrows correspond to the other. Select the structure that you think you had.





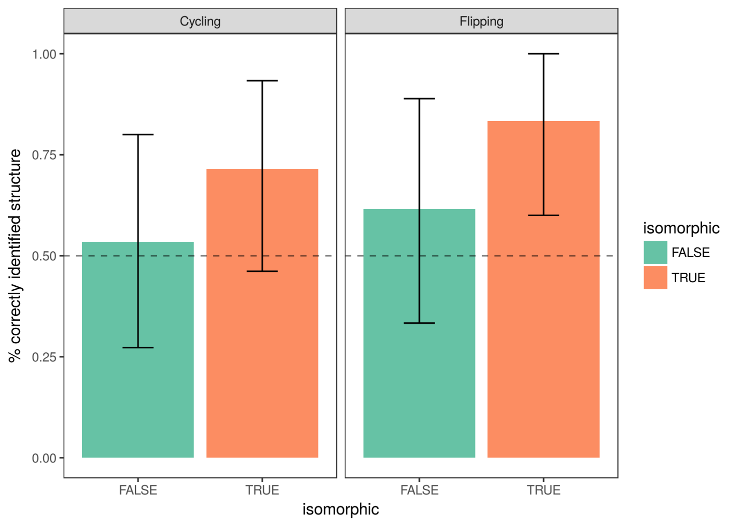
Submit Answers

Isomorphism may influence within-task abstraction



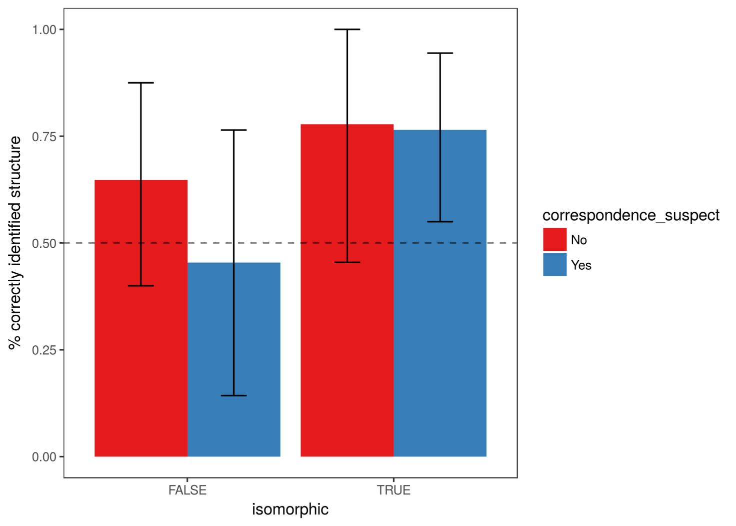
Error bars represent bootstrap 95%-confidence intervals

Isomorphism may influence within-task abstraction



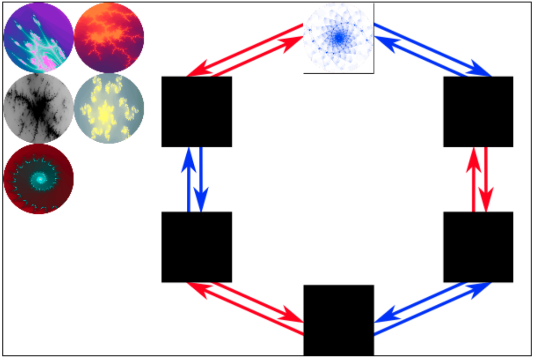
Error bars represent bootstrap 95%-confidence intervals

Isomorphism may influence within-task abstraction



Error bars represent bootstrap 95%-confidence intervals

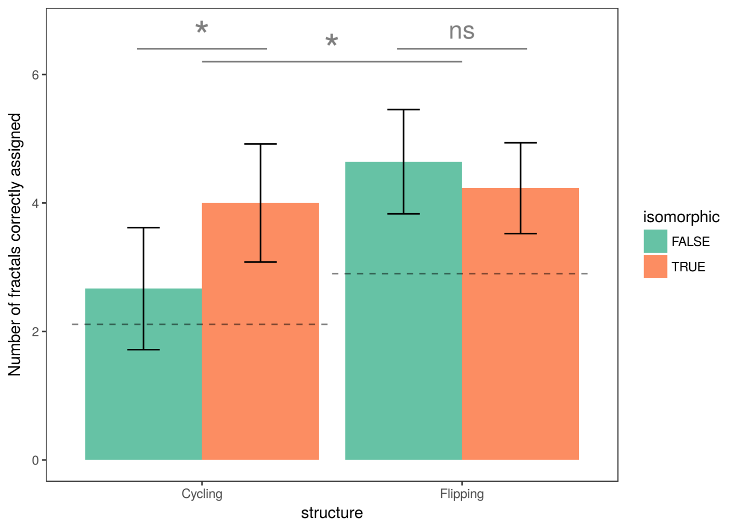
Isomorphism may influence within-task abstraction



The fractal relationships in your experiment obeyed the structure above (black squares = fractals, red arrows = acid, blue arrows = gamma ray). Drag the fractals onto the black squares that you think they map onto (one has been placed to get you started).

Submit answers

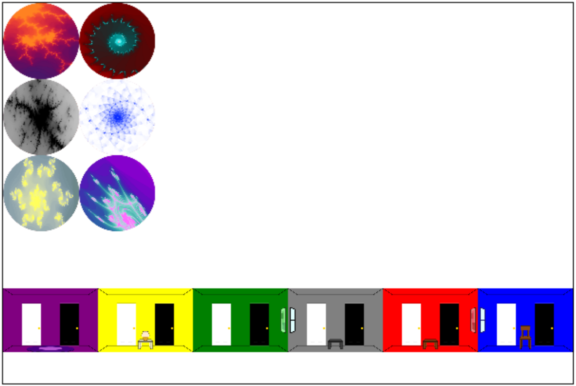
Isomorphism may influence within-task abstraction



Error bars represent bootstrap 95%-confidence intervals

Results: Explicit knowledge about relationship

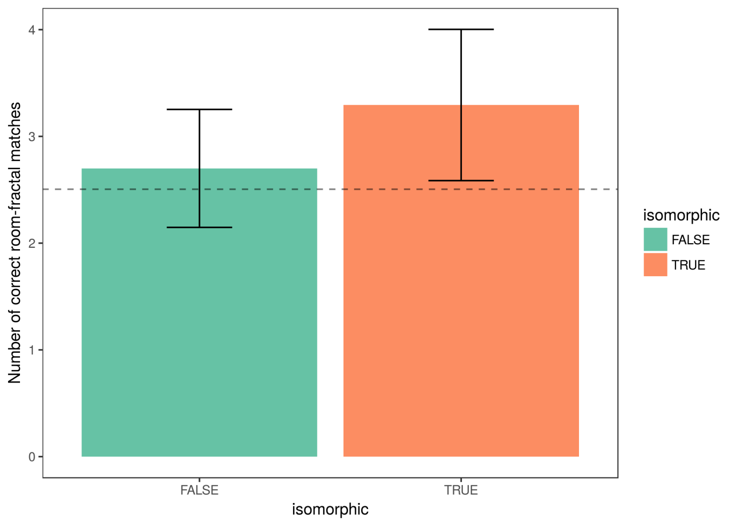
Subjects can explicitly identify analogy better than chance



Assuming the experiments were teaching you the same structure, each room in the navigation experiment corresponded to a fractal in the visual patterns experiment. Drag the fractals onto the rooms you think they corresponded to.

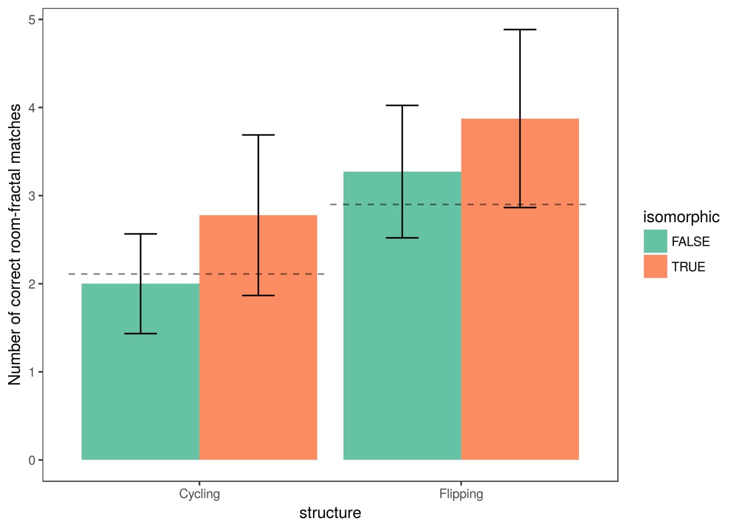
Submit answers

Subjects can explicitly identify analogy better than chance



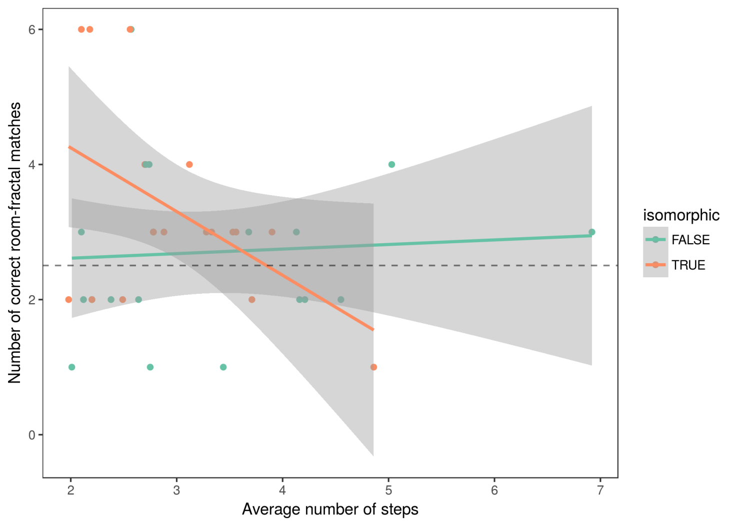
Error bars represent parametric 95%-confidence intervals

Subjects can explicitly identify analogy better than chance

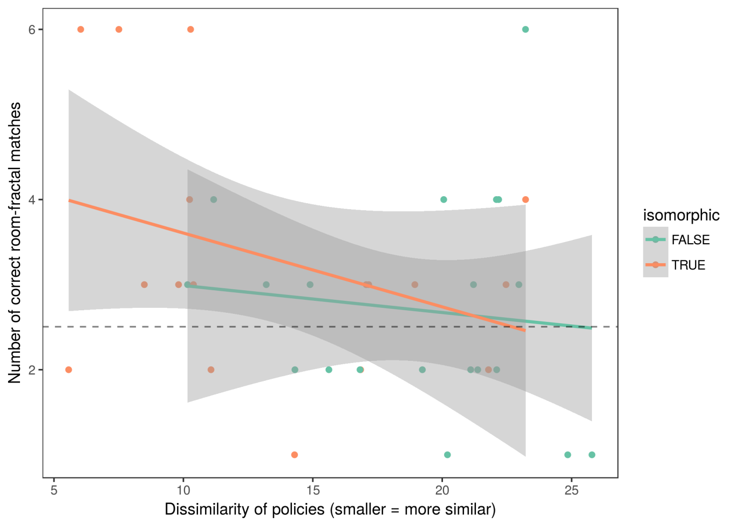


Error bars represent parametric 95%-confidence intervals

Only weakly related to performance and policy alignment

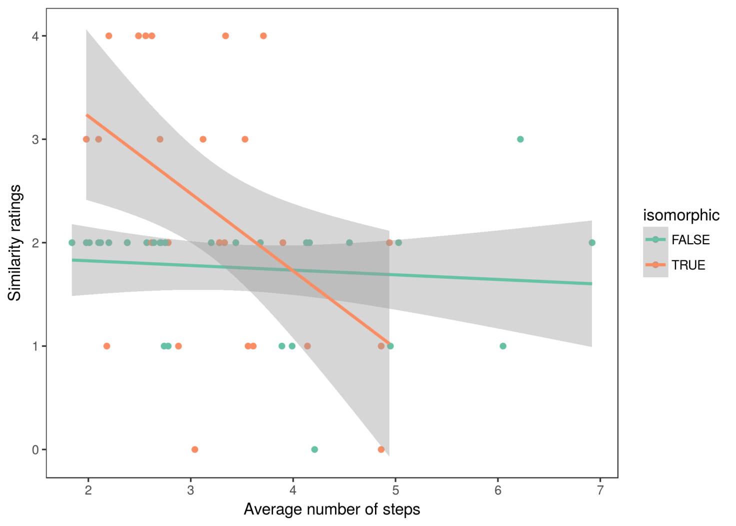


Only weakly related to performance and policy alignment



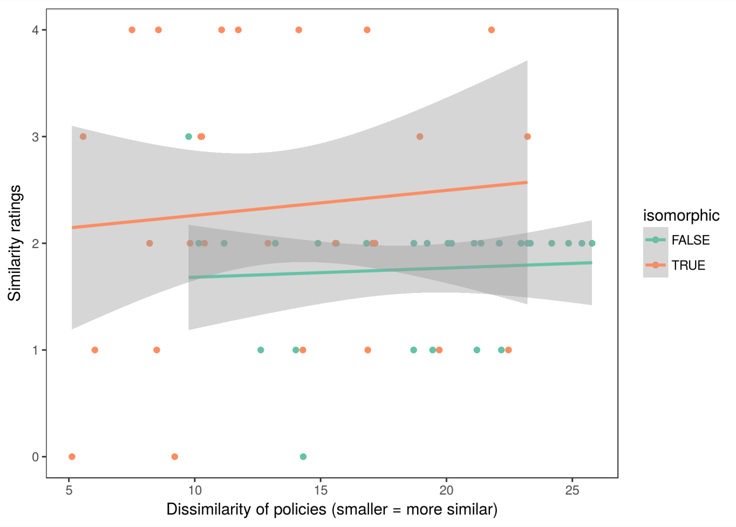
Error bars represent parametric 95%-confidence intervals

Similarity ratings



Error bars represent parametric 95%-confidence intervals

Similarity ratings



Error bars represent parametric 95%-confidence intervals

Slides TODO: but policy alignment does not seem to predict anything explicit very strongly.

Conclusions

- We see evidence of slow transfer between tasks.

Conclusions

- We see evidence of slow transfer between tasks.
- This slow transfer may be independent of consciousness.

Conclusions

- We see evidence of slow transfer between tasks.
- This slow transfer may be independent of consciousness.
- This slow transfer of knowledge affects within-task explicit knowledge.

Conclusions

- We see evidence of slow transfer between tasks.
- This slow transfer may be independent of consciousness.
- This slow transfer of knowledge affects within-task explicit knowledge.
- However, the processes for implicit and explicit mapping between tasks may be at least partially dissociable.

Next steps & possible future directions

- See whether these results hold/replicate

Next steps & possible future directions

- See whether these results hold/replicate
- Verify that direction of transfer is caused by order and not e.g. grounding

Next steps & possible future directions

- See whether these results hold/replicate
- Verify that direction of transfer is caused by order and not e.g. grounding
- Distractor tasks between

Next steps & possible future directions

- See whether these results hold/replicate
- Verify that direction of transfer is caused by order and not e.g. grounding
- Distractor tasks between
- Integrating abstractions (either taught or interrogated) earlier on one or both tasks

Next steps & possible future directions

- See whether these results hold/replicate
- Verify that direction of transfer is caused by order and not e.g. grounding
- Distractor tasks between
- Integrating abstractions (either taught or interrogated) earlier on one or both tasks
- Role of sleep in abstraction

Acknowledgements

References

Douglas K. Detterman. The Case for the Prosecution: Transfer as an Epiphenomenon. In *Transfer on Trial: Intelligence, Cognition, and Instruction*, pages 1–24. 1993.

Dedre Gentner. Why We're So Smart. In *Language in mind: Advances in the study of language and thought.*, pages 195–235. 2003.

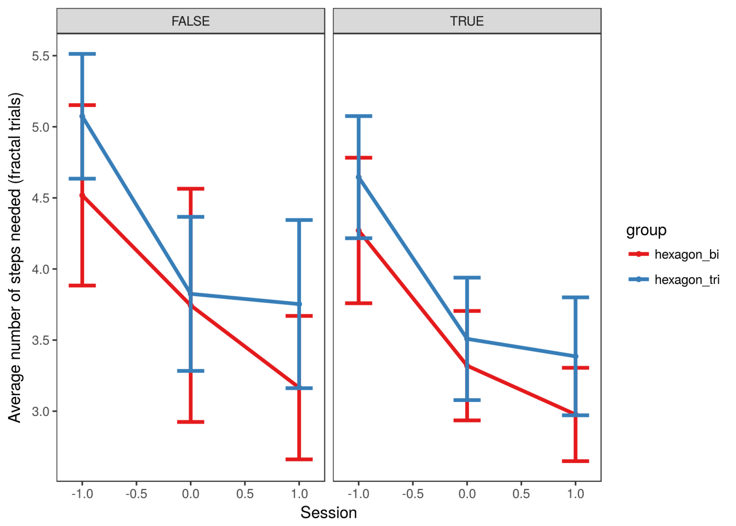
Mary L Gick and Keith J. Holyoak. Analogical Problem Solving. *Cognitive P*, 12:306–355, 1980.

Steven S. Hansen, Andrew Lampinen, Gaurav Suri, and James L. McClelland. Building on prior knowledge without building it in. *Behavioral and Brain Sciences*, 40, 2017.

Andrew Lampinen, Shaw Hsu, and James L Mcclelland. Analogies Emerge from Learning Dyamics in Neural Networks. *Proceedings of the 39th Annual Conference of the Cognitive Science Society*, pages 2512–2517, 2017.

Ben R. Newell and David R. Shanks. Unconscious influences on decision making: A critical review. *Behavioral and Brain Sciences*, 38(01):1–19, 2014. ISSN 0140-525X. doi: 10.1017/S0140525X12003214. URL `http://www.journals.cambridge.org/abstract{_}S0140525X12003214`.

Flipping is easier



Error bars represent 95%-confidence intervals

Average number of steps needed
FALSE 1
FALSE 3
FALSE 5
FALSE 7
FALSE 9
FALSE 13
FALSE 17
FALSE 18
FALSE 21
FALSE 22
FALSE 24
FALSE 25
FALSE 26
FALSE 34
FALSE 35
FALSE 37
FALSE 40
FALSE 42
FALSE 45
FALSE 47
FALSE 49
FALSE 53
FALSE 55
FALSE 56
FALSE 57
FALSE 58
FALSE 60
FALSE 61
FALSE 62
TRUE 2
TRUE 4
TRUE 6
TRUE 8
TRUE 11
TRUE 12
TRUE 14
TRUE 15
TRUE 16
TRUE 19
TRUE 23
TRUE 27
TRUE 29
TRUE 30
TRUE 32
TRUE 33
TRUE 36
TRUE 38
TRUE 39
TRUE 41
TRUE 43
TRUE 44
TRUE 46
TRUE 48
TRUE 50
TRUE 51
TRUE 52
0
5
10
15
structure
Flipping
Cycling
trial_type
door
fractal