New Experimental Evidence against the Similarity Approach to Conditionals

Dean McHugh and Tomasz Klochowicz

Institute of Logic, Language and Computation
University of Amsterdam

Università degli Studi di Padova
Monday, 21 October 2024

Plan

1. Reciprocity
   - Dynamic counterexamples to Reciprocity
   - Novel static counterexamples to Reciprocity
2. Testing the Predictions
   - Similarity Approach
   - Kratzer's Premise Semantics
   - Ciardelli et al.'s Background Semantics
   - Fine's Truthmaker Semantics
   - Aboutness Approach
3. Follow-up Experiment

Reciprocity

Also known as CSO (Nute 1980)

Let $A > C$ denote *if A, would C*.

$$\frac{A > B \qquad B > A}{(A > C) \leftrightarrow (B > C)}$$

If A and B conditionally imply each other,
they are intersubstitutable *salva veritate* in conditional antecedents.

Reciprocity

Also known as CSO (Nute 1980)

Let $A > C$ denote *if A, would C*.

$$\frac{A > B \qquad B > A}{(A > C) \leftrightarrow (B > C)}$$

If A and B conditionally imply each other,
they are intersubstitutable *salva veritate* in conditional antecedents.

A wide range of semantics of conditional validate Reciprocity,
including Stalnaker (1968) and Lewis (1973).

The ordering approach

Model
For each world w, let $\leq_w$ be (at least) a reflexive and transitive binary relation over the set of possible worlds.

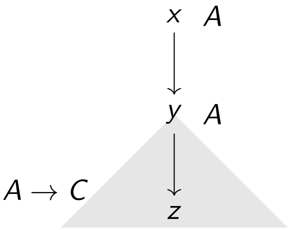
The ordering approach

Model
For each world w, let $\leq_w$ be (at least) a reflexive and transitive binary relation over the set of possible worlds.

Truth conditions (Lewis 1981, p. 230)
If A, would C is true at a world w

just in case

for every A-world x,
there is an A-world $y \leq_w x$ such that
for every world $z \leq_w y$,
if A is true at z then C is true at z.



Reciprocity is valid according to the ordering approach

Proof.

Pick any $w \in W$ and suppose that $A > B$, $B > A$ and $B > C$ are true at w. To show that $A > C$ is true at w, pick any $x \models A$. We have to show that

$$\text{there is a } y \models A \text{ such that } y \leq_w x \text{ and for all } z \leq_w y,\ z \models A \rightarrow C.$$

Since $w \models A > B$ and $x \models A$, there is a $v \models A$ such that $v \leq_w x$ and (i) for all $v' \leq_w v$, $v' \models A \rightarrow B$. Since $\leq_w$ is reflexive, $v \leq_w v$, so $v \models A \rightarrow B$. Thus $v \models B$.

Since $w \models B > A$ and $v \models B$, there is a $u \models B$ such that $u \leq_w v$ and (ii) for all $u' \leq_w u$, $u' \models B \rightarrow A$.

Since $w \models B > C$ and $u \models B$, there is a $y \models B$ such that $y \leq_w u$ and (iii) for all $z \leq_w y$, $z \models B \rightarrow C$.

Since $y \leq_w u$, by (ii), $y \models B \rightarrow A$. Then as $y \models B$, $y \models A$.

And as $y \leq_w u \leq_w v \leq_w x$, by transitivity of $\leq_w$, $y \leq_w x$.

We show that for all $z \leq_w y$, $z \models A \rightarrow C$. Pick any $z \leq_w y$. Then $z \leq_w y \leq_w u \leq_w v$, so by transitivity of $\leq_w$, $z \leq_w v$. Then by (i), $z \models A \rightarrow B$. And since $z \leq_w y$, by (iii), $z \models B \rightarrow C$. Hence $z \models A \rightarrow C$. □

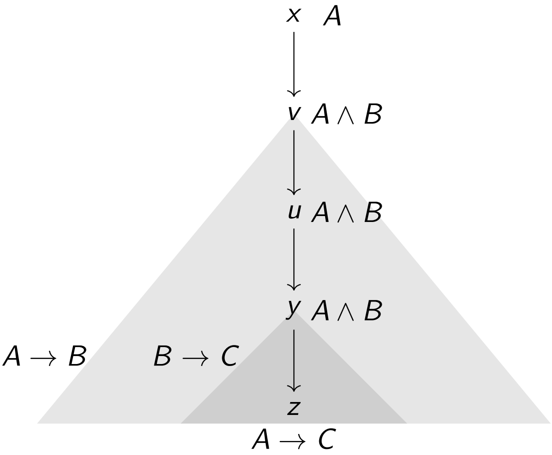


Figure: Illustrating the proof that reciprocity is valid on the ordering semantics.

Stalnaker's selection function approach validates reciprocity

> *Consider a possible world in which A is true, and which otherwise differs minimally from the actual world. 'If A, then B' is true (false) just in case B is true (false) in that possible world.*

Let W be the set of possible worlds, and $f : \wp(W) \times W \to W$ a function from propositions to worlds.

Semantics: $A > B$ is true at world w just in case B is true at $f(A, w)$.

Constraints on the selection function

1. A is true at $f(A, w)$.
2. $f(A, w)$ is the absurd world λ (the world where every proposition is true) only if there is no possible world with respect to w in which A is true.
3. If A is true in w then $f(A, w) = w$.
4. If A is true in $f(B, w)$ and B is true in $f(A, w)$, then $f(A, w) = f(B, w)$.

Constraints on the selection function

1. A is true at $f(A, w)$.
2. $f(A, w)$ is the absurd world λ (the world where every proposition is true) only if there is no possible world with respect to w in which A is true.
3. If A is true in w then $f(A, w) = w$.
4. If A is true in $f(B, w)$ and B is true in $f(A, w)$, then $f(A, w) = f(B, w)$.

Constraints on the selection function

The selection is based on an ordering of possible worlds with respect to their resemblance to the base world. If this is correct, then [(3) and (4)] must be imposed on the s-function [the selection function]. ...

These conditions on the selection function are necessary in order that this account be recognizable as an explication of the conditional.

(Stalnaker 1968, p. 36)

Plan

1. Reciprocity
   - Dynamic counterexamples to Reciprocity
   - Novel static counterexamples to Reciprocity
2. Testing the Predictions
   - Similarity Approach
   - Kratzer's Premise Semantics
   - Ciardelli et al.'s Background Semantics
   - Fine's Truthmaker Semantics
   - Aboutness Approach
3. Follow-up Experiment

A counterexample to reciprocity (Bacon 2013)

If A falls, it knocks over B, and vice versa. The balls are on sensors. If A falls while B is stationary, the light turns green. If B falls while A is stationary it turns red.

Figure: Bacon's counterexample to reciprocity.

A counterexample to reciprocity (Bacon 2013)

If A falls, it knocks over B, and vice versa. The balls are on sensors. If A falls while B is stationary, the light turns green. If B falls while A is stationary it turns red.

Figure: Bacon's counterexample to reciprocity.

(1) a. If A fell, B would fall.

A counterexample to reciprocity (Bacon 2013)

If A falls, it knocks over B, and vice versa. The balls are on sensors. If A falls while B is stationary, the light turns green. If B falls while A is stationary it turns red.

Figure: Bacon's counterexample to reciprocity.

(1) a. If A fell, B would fall.
    b. If B fell, A would fall.

A counterexample to reciprocity (Bacon 2013)

If A falls, it knocks over B, and vice versa. The balls are on sensors. If A falls while B is stationary, the light turns green. If B falls while A is stationary it turns red.

Figure: Bacon's counterexample to reciprocity.

(1) a. If A fell, B would fall.
b. If B fell, A would fall.
c. If A fell, the light would turn green.

A counterexample to reciprocity (Bacon 2013)

If A falls, it knocks over B, and vice versa. The balls are on sensors. If A falls while B is stationary, the light turns green. If B falls while A is stationary it turns red.

Figure: Bacon's counterexample to reciprocity.

(1) a. If A fell, B would fall.
b. If B fell, A would fall.
c. If A fell, the light would turn green.
d. If B fell, the light would turn red.

A counterexample to reciprocity (Bacon 2013)

If A falls, it knocks over B, and vice versa. The balls are on sensors. If A falls while B is stationary, the light turns green and stays green. If B falls while A is stationary it turns red and stays red.

(2) a. If A had fallen, B would have fallen.
b. If B had fallen, A would have fallen.
c. If A had fallen, the light would have turned green.
d. If B had fallen, the light would have turned red.

The possibility of backtracking readings

I might say to you: 'look, when A topples because it's hit by B the green light will not come on ... So if A were to topple, the green light might not come on (because B toppled first).' ...

It must be stressed, however, that our case against [Reciprocity] does not depend on the possibility of contexts in which backtracking is legitimate. We only need 1-4 to be simultaneously true in one context to complete our case against [Reciprocity], which deems them jointly inconsistent. It does not matter if there are also contexts in which some or all of 1-4 are false.

(Bacon 2013, p. 18)

Taking temporal information into account

(3)
a. If A fell [at time t], B would fall [at some time $t' \geq t$].
b. If B fell [at time t], A would fall [at some time $t' \geq t$].
c. If A fell [at time t], the light would turn green [at some time $t' \geq t$].
d. If B fell [at time t], the light would turn green [at some time $t' \geq t$].

Form:

a. $A > B'$
b. $B > A'$
c. $A > C$
d. $B > C$

Taking temporal information into account

(3) a. If A fell [at time t], B would fall [at some time $t' \geq t$].
b. If B fell [at time t], A would fall [at some time $t' \geq t$].
c. If A fell [at time t], the light would turn green [at some time $t' \geq t$].
d. If B fell [at time t], the light would turn green [at some time $t' \geq t$].

Form:

a. $A > B'$
b. $B > A'$
c. $A > C$
d. $B > C$

Not an instance of Reciprocity.

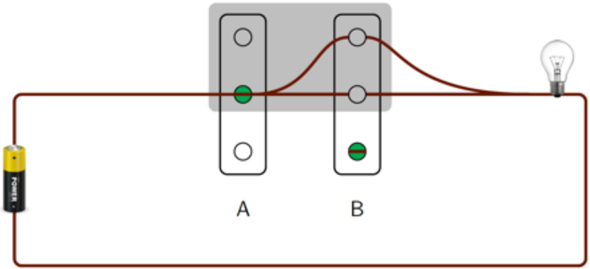
Lewis (1973) assumes a time-independent notion of overall similarity.

Bennett (1984) says: determine similarity at a particular time.

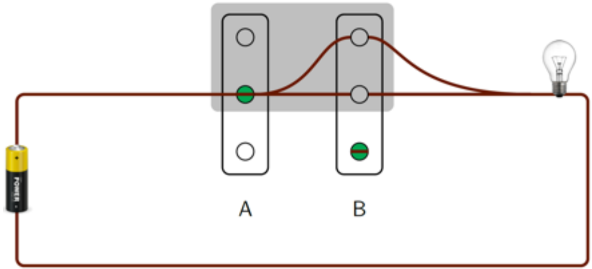
> *A counterfactual is true just in case the consequent holds at all worlds which, among those where the antecedent is true and that obey the laws of the actual world, are closest to the actual world at the time to which the antecedent pertains.*
>
> (Bennett 1984)

Plan

1. Reciprocity
   - Dynamic counterexamples to Reciprocity
   - Novel static counterexamples to Reciprocity
2. Testing the Predictions
   - Similarity Approach
   - Kratzer's Premise Semantics
   - Ciardelli et al.'s Background Semantics
   - Fine's Truthmaker Semantics
   - Aboutness Approach
3. Follow-up Experiment

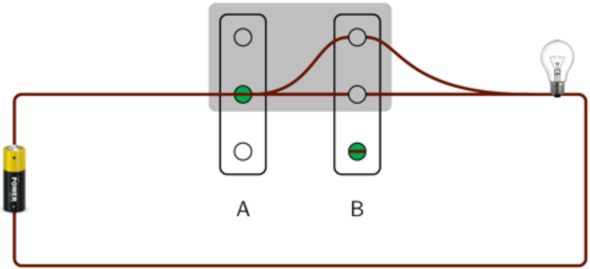


There are two switches, A and B, connected to a light. Part of the circuit is shaded grey. Each switch has three possible positions: up, in the middle, or down. The current position of the switch is indicated by a green circle.



There are two switches, A and B, connected to a light. Part of the circuit is shaded grey. Each switch has three possible positions: up, in the middle, or down. The current position of the switch is indicated by a green circle.

As the wiring indicates, the light is on just in case switch A is in the middle and switch B is either up or in the middle.

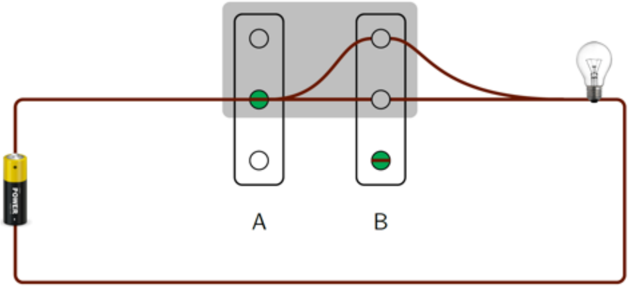


There are two switches, A and B, connected to a light. Part of the circuit is shaded grey. Each switch has three possible positions: up, in the middle, or down. The current position of the switch is indicated by a green circle.

As the wiring indicates, the light is on just in case switch A is in the middle and switch B is either up or in the middle.

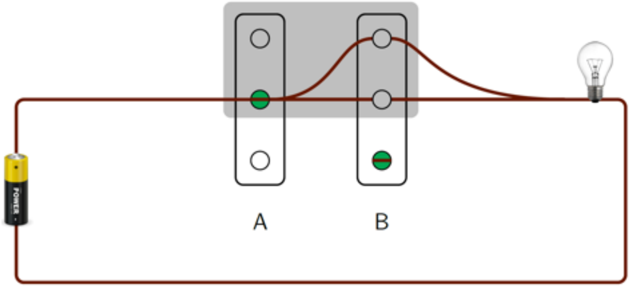
Currently, switch A is in the middle and switch B is down, so the light is off.

Premise 1



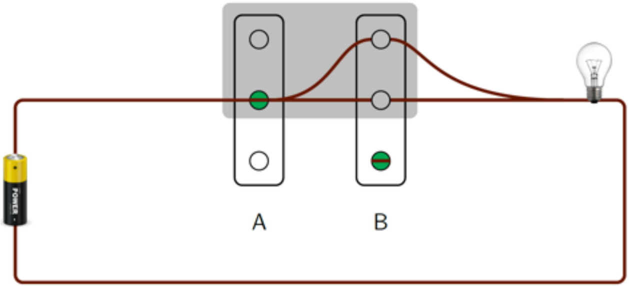
If switch B were in the shaded area,
both switches would be in the shaded area.

Premise 2



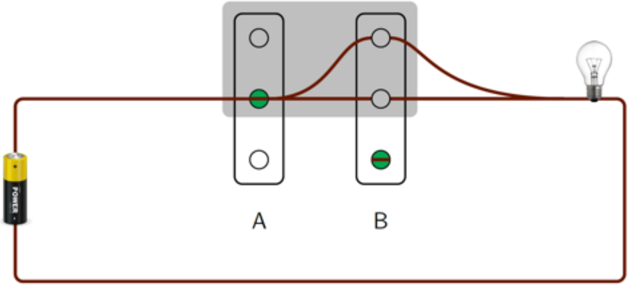
If both switches were in the shaded area,
switch B would be in the shaded area.

Premise 3

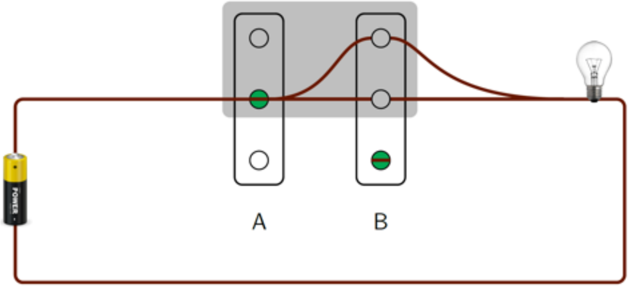


If switch B were in the shaded area,
the light would be on.

Conclusion

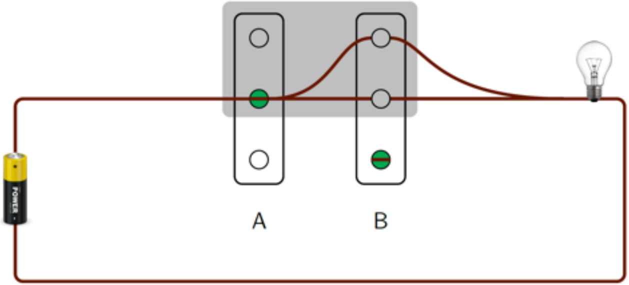


If both switches were in the shaded area,
the light would be on.



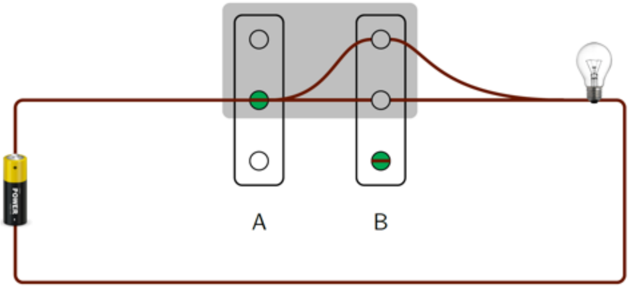
- T If both switches were in the middle, the light would be on.
- P1 If switch B were in the shaded area, both switches would be in the shaded area. *B > both*
- P2 If both switches were in the shaded area, switch B would be in the shaded area. *both > B*
- P3 If switch B were in the shaded area, the light would be on. *B > on*
- C If both switches were in the shaded area, the light would be on. *both > on*
- F If both switches were outside the shaded area, the light would be on.

True control



If both switches were in the middle,
the light would be on.

False control



If both switches were outside the shaded area, the light would be on.

Experimental design

- 80 native English speakers, recruited via Prolific.
- Following Romoli, Santorio, and Wittenberg (2022), for each sentence we asked whether it is true, false, or indeterminate.
  - If indeterminate: follow up whether they strongly feel that there is no correct answer or just do not know.
  - We excluded the latter responses from the analysis.

Experimental design

- 80 native English speakers, recruited via Prolific.
- Following Romoli, Santorio, and Wittenberg (2022), for each sentence we asked whether it is true, false, or indeterminate.
    - If indeterminate: follow up whether they strongly feel that there is no correct answer or just do not know.
    - We excluded the latter responses from the analysis.
- We collected reaction times as a measure of diffuculty of processing.

The experiment is available at
`www.tklochowicz.com/experiment_reciprocity`

Candy scenario: conclusion

Alice likes strawberry-flavoured candy and does not like any other kind of candy. Bob likes all fruit-flavoured candy, but does not like any other kind of candy.

If both children's candy were fruit-flavoured, both children would be happy with their candy.

True | Indeterminate | False

Instance of reciprocity from the candy scenario

Context: Alice likes strawberry-flavoured candy and no other flavours, while Bob likes all fruit-flavoured candy and no other flavours. The teacher gave Alice got a strawberry-flavoured candy and Bob a mint. So Alice was happy with her candy and Bob was not.

(P1) If Bob's candy had been fruit-flavoured, both of the children's candy would have been fruit-flavoured.

(P2) If both of the children's candy had been fruit-flavoured, Bob's candy would have been fruit-flavoured.

(P3) If Bob's candy had been fruit-flavoured, both of the children would have been happy with their candy.

(C) If both of the children's candy had been fruit-flavoured, both of the children would have been happy with their candy.

Experimental design

We tested 3 scenarios. In each scenario the participant answered:

- 4 training items
- 7 filler items
- 2 controls
- 3 premises and the Conclusion

Experimental design

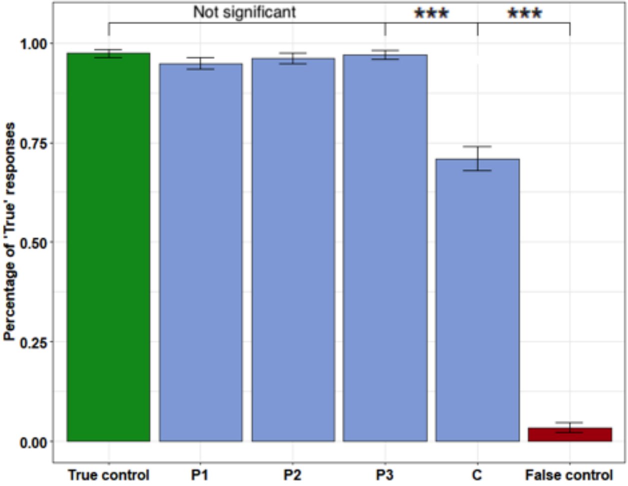
We tested 3 scenarios. In each scenario the participant answered:

- 4 training items
- 7 filler items
- 2 controls
- 3 premises and the Conclusion

Participants understood the scenarios well:

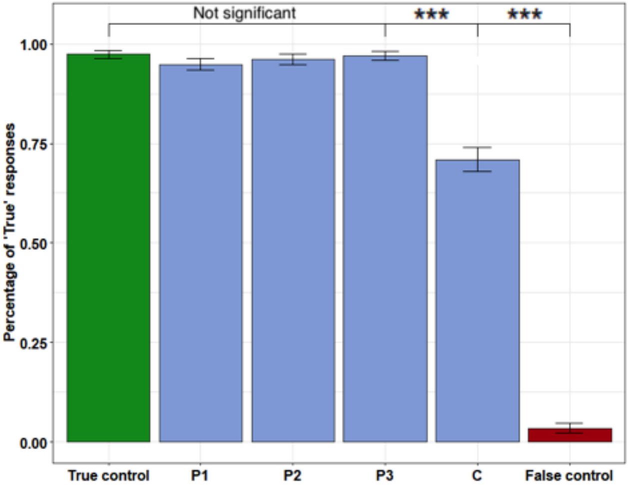
- Mean accuracy of 89% on the filler items
- We excluded two participants whose error rates on the fillers were above 30%.

Results



Not significant
***
***
Percentage of 'True' responses
1.00
0.75
0.50
0.25
0.00
True control
P1
P2
P3
C
False control

Results

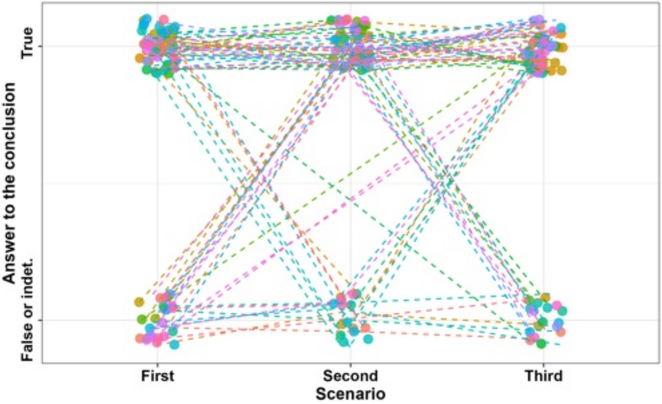


In 28% of cases where a participant accepted all of the premises P1–P3 they rejected the conclusion C.

Sentence	True	Indeterminate	False	Not sure
True control	232	0	8	0
Premise 1	225	0	11	4
Premise 2	227	2	8	3
Premise 3	225	2	12	1
Conclusion	163	26	43	8
False control	9	0	229	2

Table: Responses from all three scenarios.

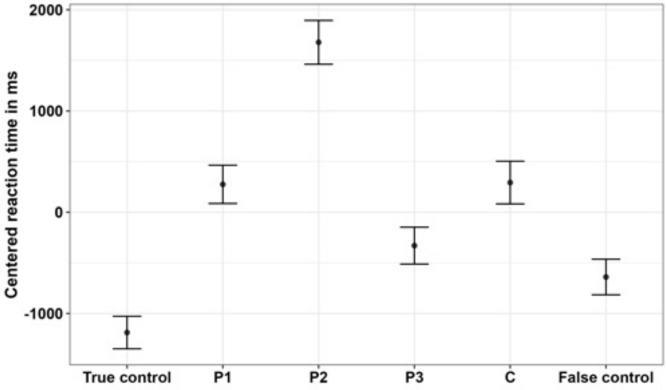
Strategy changes



Half (39/78) **of the participants accepted the premises and rejected the conclusion** in at least once in a scenario where they additionally also answered all the controls correctly.

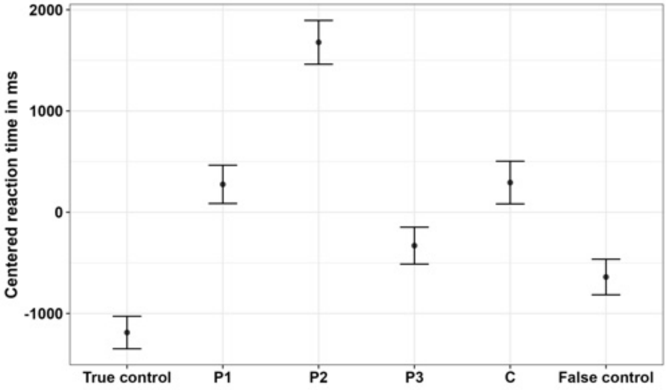
Reaction Times

Mean reaction time: 6.8 seconds and a standard deviation of 4.7 seconds.



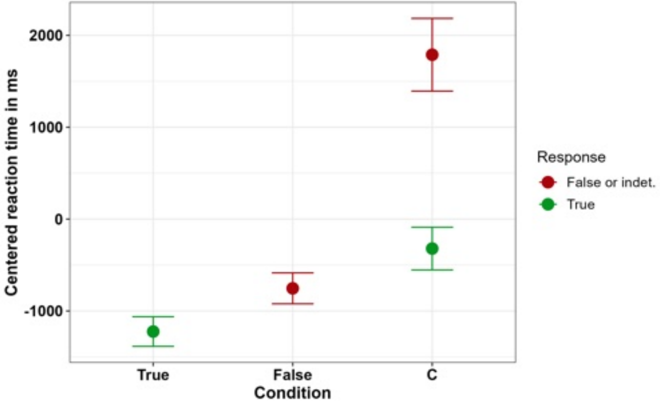
Reaction Times

Mean reaction time: 6.8 seconds and a standard deviation of 4.7 seconds.



(P2) If both of the children's candy had been fruit-flavoured, Bob's candy would have been fruit-flavoured.

Reaction times to the conclusion



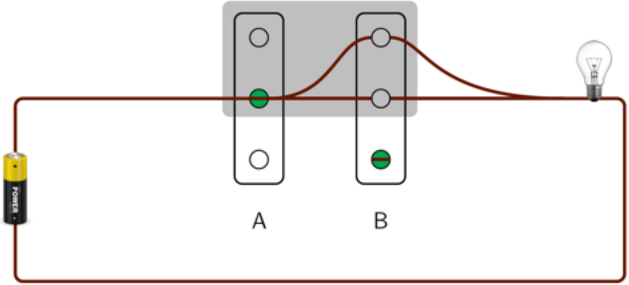
Participants took on average 1.9 seconds longer to answer *false* or *indeterminate* than to answer *true* to the Conclusion trials ($p < 0.001$).

Taking stock

- Reciprocity is not a valid inference.
- The conclusion (C) is usually evaluated as true (70% of the time).

Taking stock

- Reciprocity is not a valid inference.
- The conclusion (C) is usually evaluated as true (70% of the time).
- If participants answer *True*, they only consider moving switch B.
- Since switch A satisfies the condition of being in the shaded area, imagining it in a different position may be more costly.



If both switches were in the shaded area,
the light would be on.

1. Reciprocity
   - Dynamic counterexamples to Reciprocity
   - Novel static counterexamples to Reciprocity
2. Testing the Predictions
   - Similarity Approach
   - Kratzer's Premise Semantics
   - Ciardelli et al.'s Background Semantics
   - Fine's Truthmaker Semantics
   - Aboutness Approach
3. Follow-up Experiment

Plan

The similarity approach

1. Reciprocity is a valid inference in ordering-based approaches, regardless how we interpret the similarity order.
2. Our experimental results show that Reciprocity is not a valid inference in natural language.

The similarity approach

1. Reciprocity is a valid inference in ordering-based approaches, regardless how we interpret the similarity order.
2. Our experimental results show that Reciprocity is not a valid inference in natural language.
3. It is impossible to construct a model in Stalnaker's or Lewis's semantics in which the premises are true and the conclusion false.
4. A significant number of our participants reasoned in this way (half of them at least once).

Plan

Lumping semantics (Kratzer 1989, 2012)

> *The truth of counterfactuals depends on everything which is the case in the world under consideration: in assessing them, we have to consider all the possibilities of adding as many facts to the antecedent as consistency permits. If the consequent follows from every such possibility, then (and only then), the whole counterfactual is true.*
>
> *(Kratzer 1981, p. 201)*

Plan

Background semantics (Ciardelli, Zhang, and Champollion 2018)

Contributing to the falsity

A fact f contributes to the falsity of a proposition A at world w just in case there is some set of facts F at w that is consistent with A, but $F \cup \{f\}$ is not.

Background semantics (Ciardelli, Zhang, and Champollion 2018)

Contributing to the falsity

A fact f contributes to the falsity of a proposition A at world w just in case there is some set of facts F at w that is consistent with A, but $F \cup \{f\}$ is not.

1. A fact is only allowed to change when it contributes to the falsity of the antecedent in the actual world.

Background semantics (Ciardelli, Zhang, and Champollion 2018)

Contributing to the falsity

A fact f contributes to the falsity of a proposition A at world w just in case there is some set of facts F at w that is consistent with A, but $F \cup \{f\}$ is not.

1. A fact is only allowed to change when it contributes to the falsity of the antecedent in the actual world.
2. Switch A being in the middle does not contribute to the falsity of both switches being in the shaded area.

Background semantics (Ciardelli, Zhang, and Champollion 2018)

Contributing to the falsity

A fact f contributes to the falsity of a proposition A at world w just in case there is some set of facts F at w that is consistent with A, but $F \cup \{f\}$ is not.

1. A fact is only allowed to change when it contributes to the falsity of the antecedent in the actual world.
2. Switch A being in the middle does not contribute to the falsity of both switches being in the shaded area.
3. Background semantics predicts: Keep switch A in the middle.

Background semantics (Ciardelli, Zhang, and Champollion 2018)

Contributing to the falsity

A fact f contributes to the falsity of a proposition A at world w just in case there is some set of facts F at w that is consistent with A, but $F \cup \{f\}$ is not.

1. A fact is only allowed to change when it contributes to the falsity of the antecedent in the actual world.
2. Switch A being in the middle does not contribute to the falsity of both switches being in the shaded area.
3. Background semantics predicts: Keep switch A in the middle.

Background semantics predicts the conclusion to be true.

C If both switches were in the shaded area, the light would be on.

Plan

Truthmaker semantics of conditionals (Fine 2012)

Universal Realisability of the Antecedent (Fine 2012, p. 236)

A conditional is true only if it is true for any way in which its antecedent might be true.

Truthmaker semantics of conditionals (Fine 2012)

Universal Realisability of the Antecedent (Fine 2012, p. 236)

A conditional is true only if it is true for any way in which its antecedent might be true.

Assumptions:

- The state of A being up exactly verifies *switch A is in the shaded area*.
- The state of B being up exactly verifies *switch B is in the shaded area*.

Truthmaker semantics (Fine 2012)

Analyse "both switches are in the shaded area" as a conjunction:

"switch A is in the shaded area and switch B is in the shaded area".

Truthmaker semantics (Fine 2012)

Analyse "both switches are in the shaded area" as a conjunction:

"switch A is in the shaded area and switch B is in the shaded area".

Exact verification

A state s exactly verifies $A \wedge B$ just in case s is the fusion of states t and t' where t exactly verifies A and t' exactly verifies B (Fine 2017).

The state of both switches being up exactly verifies that both switches are in the shaded area. And if both switches are up, the light is off.

Truthmaker semantics (Fine 2012)

Analyse "both switches are in the shaded area" as a conjunction:
"switch A is in the shaded area and switch B is in the shaded area".

Exact verification

A state s exactly verifies $A \wedge B$ just in case s is the fusion of states t and t' where t exactly verifies A and t' exactly verifies B (Fine 2017).

The state of both switches being up exactly verifies that both switches are in the shaded area. And if both switches are up, the light is off.

The problem for Fine

Fine (2012) predicts the conclusion to be false.
Problem: a majority of our participants judged it true.

Plan

An aboutness approach to conditionals (McHugh 2023)

1. Pick a time at which to imagine changing the change.
2. Allow the part of the world the antecedent is about at that time to vary.
3. Play forward the laws.
4. Stick on the actual past.
5. Restrict to worlds where the antecedent holds.
6. Check whether the consequent holds at the resulting world(s).

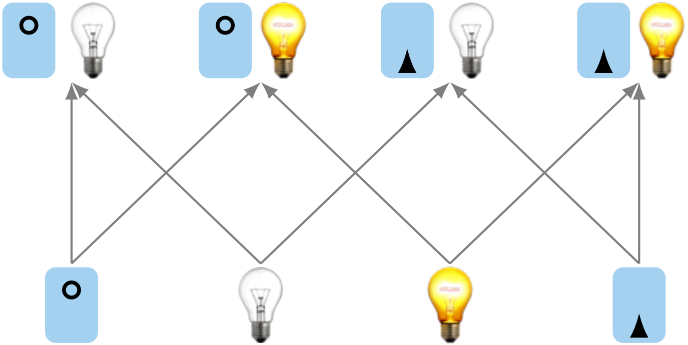


Figure: A state space of the switch and light.

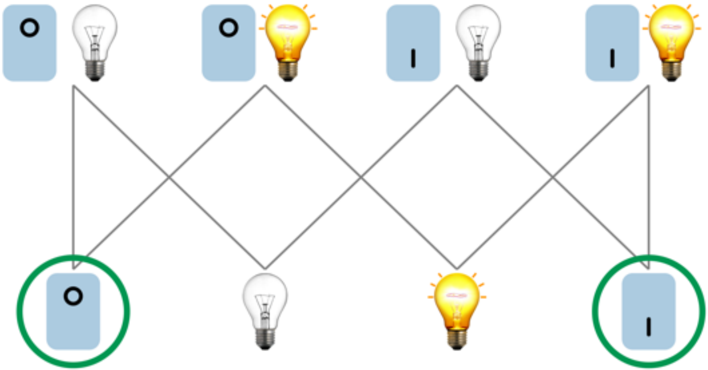


Figure: The states that "the switch is up" is about.

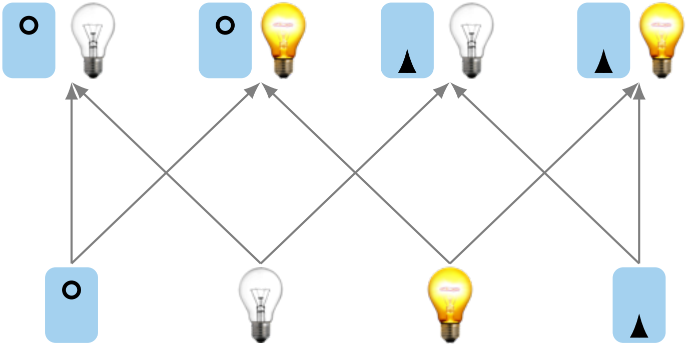
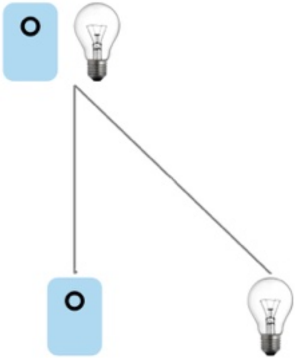
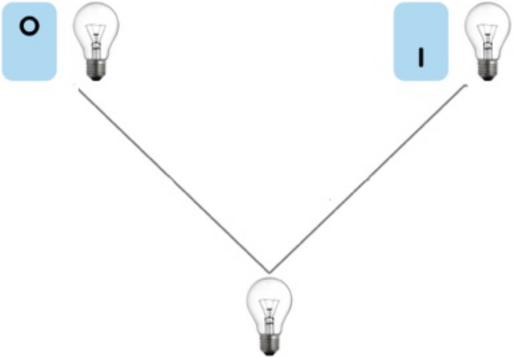


Figure: A state space of the switch and light.




O
I

Some sequences of states are lawful (or, nomically possible) and others are not.

Figure: A lawful sequence.

Some sequences of states are lawful (or, nomically possible) and others are not.

Figure: A lawful sequence.

Figure: An unlawful sequence.

- **The foreground:** the set of states A is about.
- **The background:** the set of states that do not overlap a state in the foreground.

- **The foreground:** the set of states A is about.
- **The background:** the set of states that do not overlap a state in the foreground.

Ceteris paribus

- The background is the *ceteris*, the ‘all else’ in ‘all else being equal’
- *Paribus* means having the *ceteris* as part

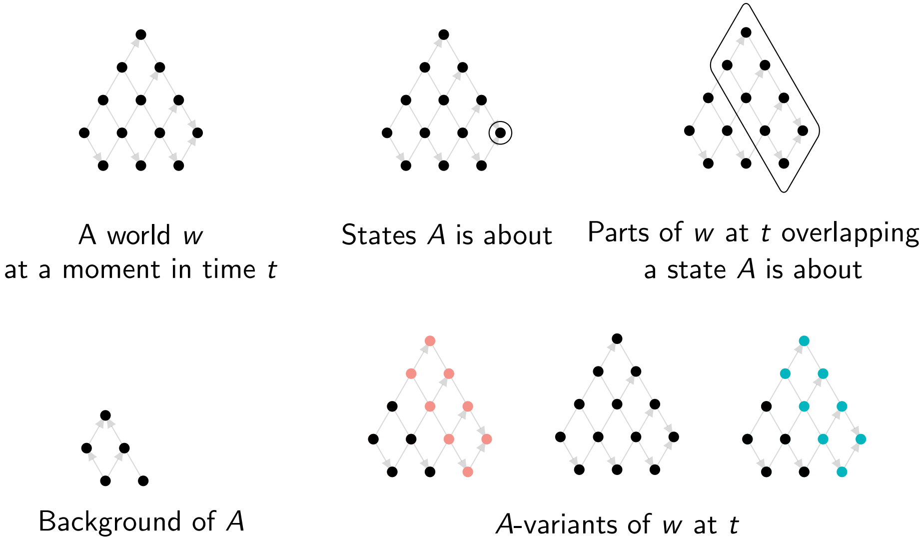


Figure: Steps to construct the A-variants of a world at a moment in time.

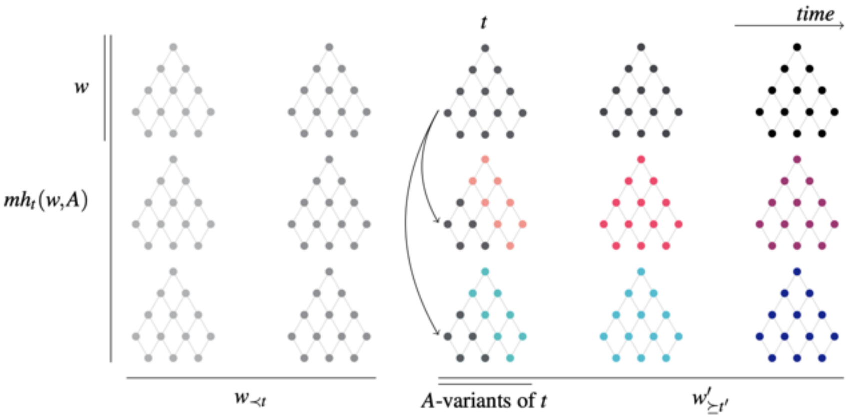


Figure: Constructing the modal horizon.

Definition (Nomic aboutness model)

Where S is a set and $\leq$ a binary relation on S, define

$$Sit := S \times I, \text{ where } I \text{ is an arbitrary label set,}$$
$$M := \{t_i \in Sit : t \leq u \textit{ implies } t = u \textit{ for all } u \in S\},$$
$$W := \{(M', \preceq) : M' \subseteq M, \preceq \text{ is a linear order}\}.$$

Definition (The modal horizon)

For any sentence A, moment $t \in M$ and world $w \in W$, define

$$mh_{P,t}(w, A) := \{w_{\prec t} \frown w'_{\succeq t'} : t' \textit{ is an } A\textit{-variant of } t, t' \in w' \text{ and } w' \in P\}.$$

(4) Where P is the set of nomically possible worlds, t the intervention time, and s the selection function,

$$A \gg C \text{ is true at } w \quad \text{iff} \quad mh_{P,t}(w, A) \cap |A| \subseteq |C|$$
$$A > C \text{ is true at } w \quad \text{iff} \quad s\big(w, mh_{P,t}(w, A) \cap |A|\big) \in |C|$$

Analyses of aboutness

Exact verification: the intuitive idea (Fine 2017)

A state *exactly verifies* a sentence just in case the state obtaining is wholly relevant to the sentence being true.
A state *exactly falsifies* a sentence just in case the state obtaining is wholly relevant to the sentence being false.

Analyses of aboutness

Exact verification: the intuitive idea (Fine 2017)

A state *exactly verifies* a sentence just in case the state obtaining is wholly relevant to the sentence being true.
A state *exactly falsifies* a sentence just in case the state obtaining is wholly relevant to the sentence being false.

- **Truthmaker view:** A sentence is about its exact verifiers and falsifiers.
- **Subject matter view:** A sentence is about the exact verifiers and falsifiers of its atomics.

The truthmaker view

Both switches are in the shaded area.

The truthmaker view

Both switches are in the shaded area.

The sentence is currently false, so we look to its exact falsifies.

A state exactly falsifies $A \wedge B$ just in case it exactly falsifies A or it exactly falsifies B.

The truthmaker view

Both switches are in the shaded area.

The sentence is currently false, so we look to its exact falsifies.

A state exactly falsifies $A \wedge B$ just in case it exactly falsifies A or it exactly falsifies B.

- The state of switch A being in the middle does not exactly verify that both switches are in the shaded area.
- The state of switch A being in the middle does not exactly falsify that both switches are in the shaded area.
- The state of switch A being down exactly falsifies that both switches are in the shaded are.

The truthmaker view

Both switches are in the shaded area.

The sentence is currently false, so we look to its exact falsifies.

A state exactly falsifies $A \wedge B$ just in case it exactly falsifies A or it exactly falsifies B.

- The state of switch A being in the middle does not exactly verify that both switches are in the shaded area.
- The state of switch A being in the middle does not exactly falsify that both switches are in the shaded area.
- The state of switch A being down exactly falsifies that both switches are in the shaded are.

Truthmaker view: "Both switches are in the shaded area" is about switch B being down, but not about A being in the middle.

The truthmaker view

Both switches are in the shaded area.

The sentence is currently false, so we look to its exact falsifies.

A state exactly falsifies $A \wedge B$ just in case it exactly falsifies A or it exactly falsifies B.

- The state of switch A being in the middle does not exactly verify that both switches are in the shaded area.
- The state of switch A being in the middle does not exactly falsify that both switches are in the shaded area.
- The state of switch A being down exactly falsifies that both switches are in the shaded are.

Truthmaker view: "Both switches are in the shaded area" is about switch B being down, but not about A being in the middle.
We vary B but not A, so this view of aboutness predicts the Conclusion to be true.

The subject matter view

Both switches are in the shaded area.

The subject matter view

Both switches are in the shaded area.

Atomics:

- Switch A is in the shaded area
- Switch B is in the shaded area

The subject matter view

Both switches are in the shaded area.

Atomics:

- Switch A is in the shaded area
- Switch B is in the shaded area

- The state of switch A being in the middle exactly verifies that switch A is in the shaded area.
- The state of switch B being in down exactly falsifies that switch B is in the shaded area.

The subject matter view

Both switches are in the shaded area.

Atomics:

- Switch A is in the shaded area
- Switch B is in the shaded area

- The state of switch A being in the middle exactly verifies that switch A is in the shaded area.
- The state of switch B being in down exactly falsifies that switch B is in the shaded area.

Subject matter view: "Both switches are in the shaded area" is about the state of A being in the middle and about the state of switch B being down.

The subject matter view

Both switches are in the shaded area.

Atomics:

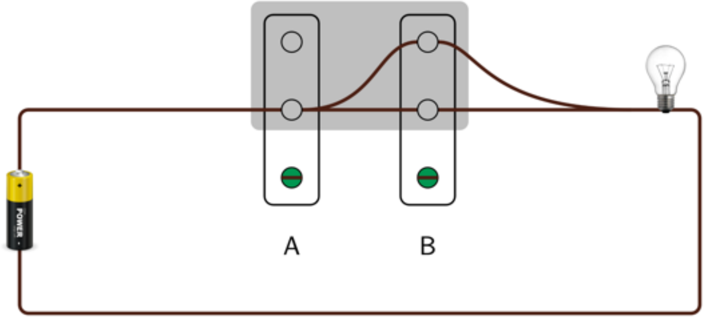
- Switch A is in the shaded area
- Switch B is in the shaded area

- The state of switch A being in the middle exactly verifies that switch A is in the shaded area.
- The state of switch B being in down exactly falsifies that switch B is in the shaded area.

Subject matter view: "Both switches are in the shaded area" is about the state of A being in the middle and about the state of switch B being down.

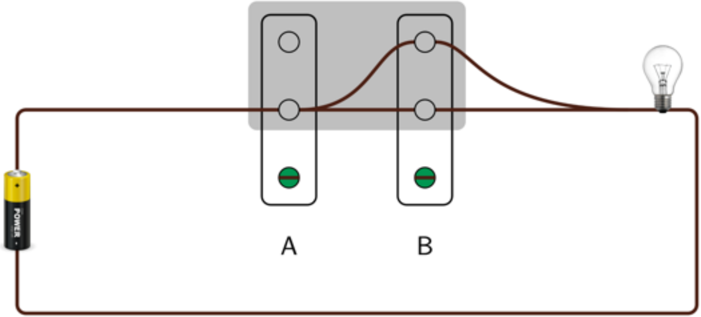
We vary both A and B, so this view of aboutness predicts the Conclusion to be false.

The follow up



C: If both switches were in the shaded area,
the light would be on.

The follow up



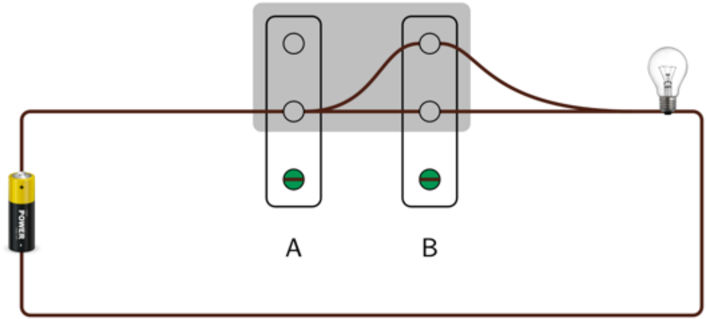
C: If both switches were in the shaded area,
the light would be on.

- Prediction: C should not be true under both analyses of aboutness.

Follow-up: Design

- Two scenarios in two blocks in random order.
- Each participant saw one conclusion of each type (each from a different scenario).
- Only the conclusion, the controls and the fillers (some premises are false in the follow-up).

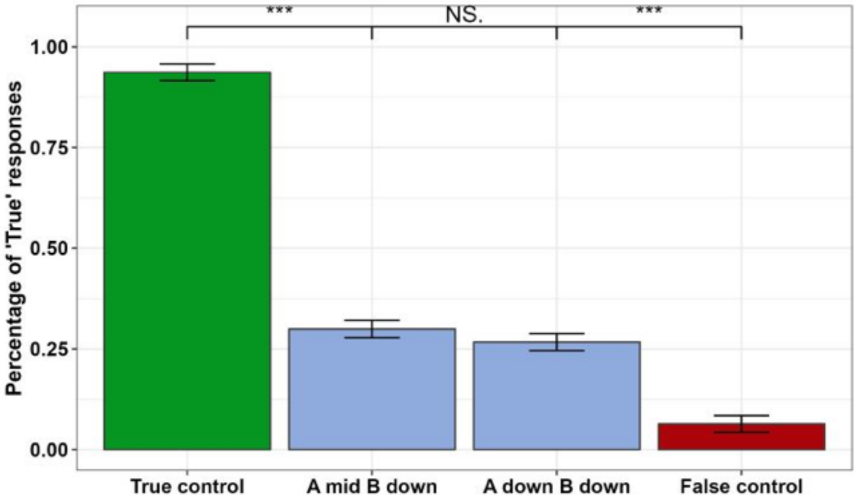
Follow-up: Design



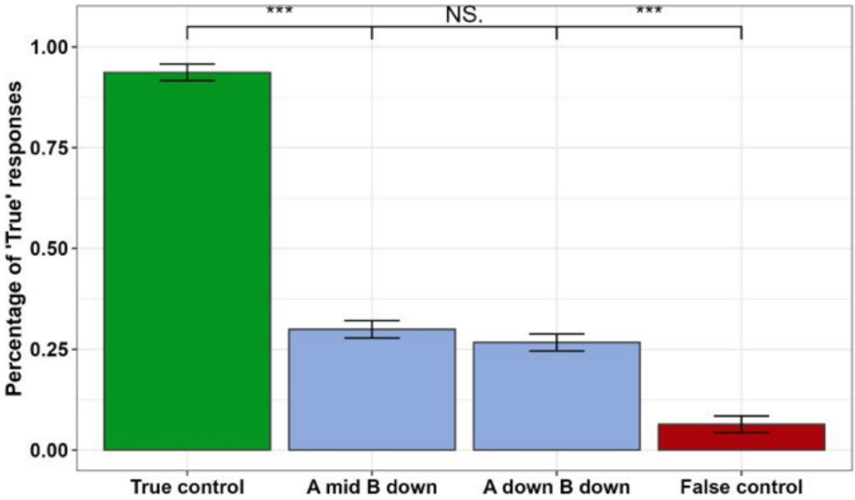
F If both switches were in the middle, the light would be off.

T If both switches were in the middle, the light would be on.

Follow-up: Results



Follow-up: Results



The only interesting effect: The *both down* conclusion is judged false significantly more often (21% acceptance) when it appears second.

Counterfactual statements can be undefined

- Counterfactual antecedents can raise many possibilities
- Consequent true in all of them: conditional is true
- Consequent false in all of them: conditional is false
- The possibilities disagree on the consequent: conditional receives a mixed response (Ramotowska et al. 2023)

Counterfactual statements can be undefined

- Counterfactual antecedents can raise many possibilities
- Consequent true in all of them: conditional is true
- Consequent false in all of them: conditional is false
- The possibilities disagree on the consequent: conditional receives a mixed response (Ramotowska et al. 2023)

Half of the tickets that have been bought will win a prize (and half will not):

(5) If ticket #37 was bought, it would win a prize.

Conclusions

1. A view variety of theories of conditionals validate Reciprocity
   - Including Stalnaker and Lewis's theories
2. Previous counterexamples to Reciprocity are inconclusive
3. Our experiment presents a novel challenge to the validity of Reciprocity
4. Our theories need some flexibility to predict the intermediate status of the Conclusion

Thank you!

References I

- Bacon, Andrew (2013). In defence of a naıve conditional epistemology. *Manuscript*.
- Bennett, Jonathan (1984). Counterfactuals and temporal direction. *The Philosophical Review* 93.1, pp. 57–91. DOI: 10.2307/2184413.
- Ciardelli, Ivano, Linmin Zhang, and Lucas Champollion (2018). Two switches in the theory of counterfactuals. *Linguistics and Philosophy* 41.6, pp. 577–621. DOI: 10.1007/s10988-018-9232-4.
- Fine, Kit (2012). Counterfactuals without possible worlds. *Journal of Philosophy* 109.3, pp. 221–246. DOI: 10.5840/jphil201210938.
- — (2017). Truthmaker Semantics. *A Companion to the Philosophy of Language*. Wiley Blackwell. Chap. 22, pp. 556–577. DOI: 10.1002/9781118972090.ch22.
- Kratzer, Angelika (1981). Partition and revision: The semantics of counterfactuals. *Journal of Philosophical Logic* 10.2, pp. 201–216. DOI: 10.1007/BF00248849.

References II

Kratzer, Angelika (1989). An investigation of the lumps of thought. *Linguistics and philosophy*, pp. 607–653. DOI: 10.1007/BF00627775.

— (2012). *Modals and conditionals*. Oxford University Press. DOI: 10.1093/acprof:oso/9780199234684.001.0001.

Lewis, David (1973). *Counterfactuals*. Wiley-Blackwell.

— (1981). Ordering semantics and premise semantics for counterfactuals. *Journal of philosophical logic*, pp. 217–234. DOI: 10.1007/BF00248850.

McHugh, Dean (2023). Causation and Modality: Models and Meanings. PhD thesis. University of Amsterdam. URL: https://eprints.illc.uva.nl/id/eprint/2243.

Nute, Donald (1980). *Topics in conditional logic*. Springer.

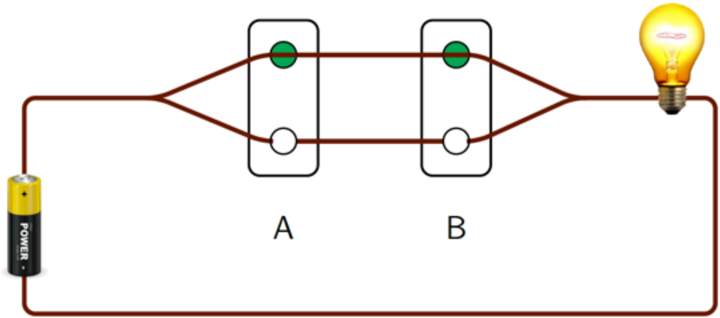
Ramotowska, Sonia et al. (2023). *Undefined counterfactuals: Experimental evidence for selection semantics*.

References III

Romoli, Jacopo, Paolo Santorio, and Eva Wittenberg (2022). Alternatives in Counterfactuals: What Is Right and What Is Not. *Journal of Semantics* 39.2, pp. 213–260. DOI: 10.1093/jos/ffab023.

Stalnaker, Robert (1968). A theory of conditionals. *Ifs*. Ed. by William L. Harper, Robert Stalnaker, and Glenn Pearce. Springer, pp. 41–55. DOI: 10.1007/978-94-009-9117-0_2.

Ciardelli, Zhang, and Champollion 2018's example.



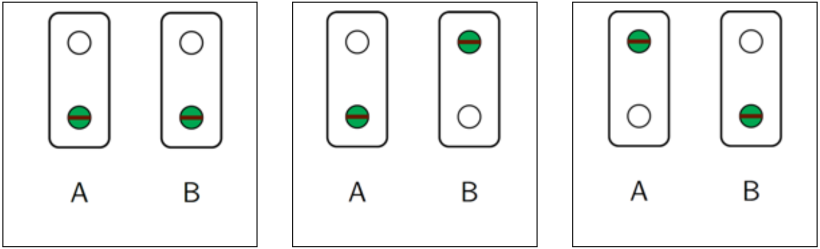
- If switch A or switch B was down, the light would be off.
- If switch A and switch B were not both up, the light would be off.

Romoli, Santorio, and Wittenberg 2022's criticism

The effect is linked to the presence of overt negation in *not both up*.

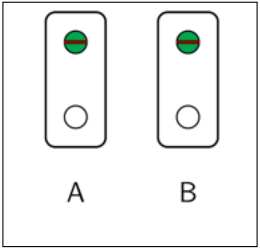
Getting rid of negation

In the following configurations, the switches follow the rules:

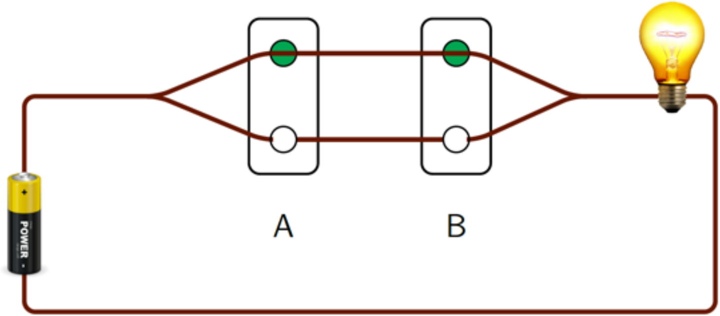


Getting rid of negation

In the following configurations, the switches are against the rules:



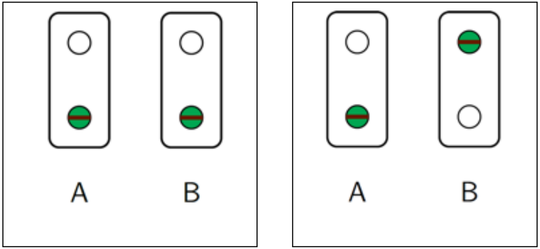
Two switches without overt negation.



- If switch A or switch B was down, the light would be off.
- If the switches **were following the rules**, the light would be off.

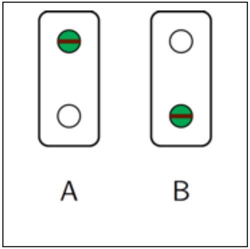
Controlling for alternatives

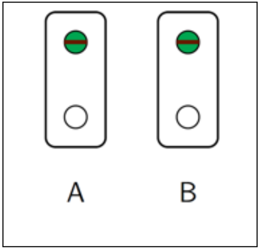
In the following configurations, switch A follows the rules:



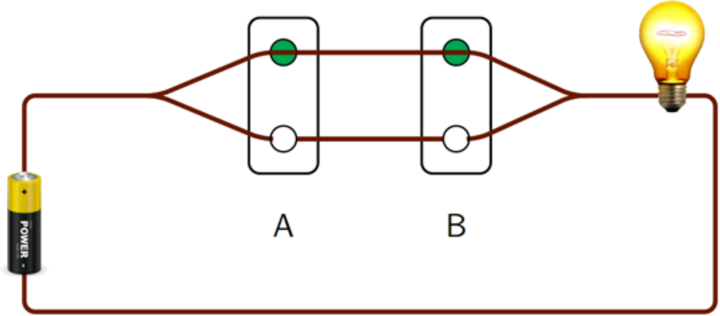
Controlling for alternatives

In the following configurations, switch A is against the rules:





Switch A without overt negation



- If switch A **was following the rules**, the light would be off.
- If switch B **was following the rules**, the light would be off.

Substitution in conditional antecedents

Substitution:

$$\frac{A \equiv B}{(A > C) \leftrightarrow (B > C)}$$

Substitution is a stronger inference than reciprocity (since $A \rightarrow B$ implies $A > B$).
Any counterexample to substitution also invalidates reciprocity.

Testing substitution

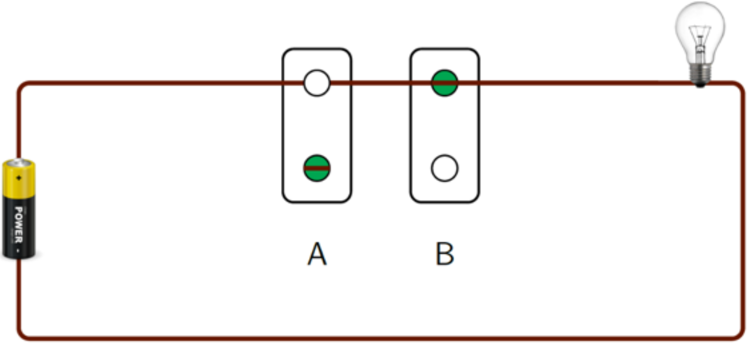
In the building where the switches are located, there are rules governing the position of switch A. Switch A can be positioned according to the rules or against the rules.

In the following configurations, **switch A is positioned according to the rules**:

In the following configurations, **switch A is positioned against to the rules**:

OK

Testing substitution



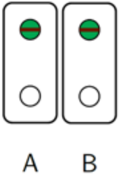
If switch A were positioned according to the rules, the light would be on.

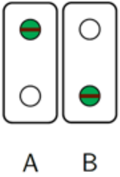
True | Indeterminate | False

Testing substitution

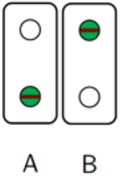
In the building where the switches are located, there are rules governing the switches. The switches can be positioned according to the rules or against the rules.

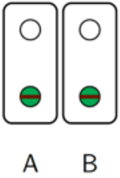
In the following configurations, **the switches are positioned according to the rules**:





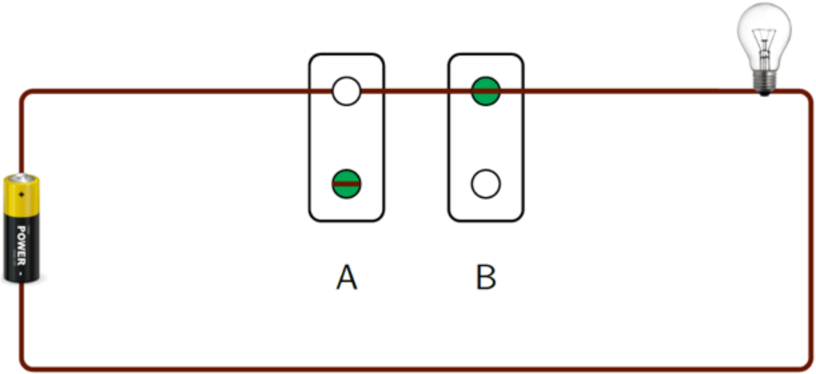
In the following configurations, **the switches are positioned against to the rules**:





OK

Testing substitution



If the switches were positioned according to the rules, the light would be on.

True | Indeterminate | False