What Do We Learn from a Conditional?

Eidon Vale*

Abstract

Suppose one learns that a reactor will overheat if it is restarted. Does one thereby acquire evidence that the reactor will not be restarted? Competing rules for learning conditionals give different answers. I argue that no answer follows from conditional content alone. A trivalent semantics distinguishes cases in which a conditional succeeds, fails, or does not apply. But this partition does not identify a unique event on which a Bayesian should condition, or determine how a report bears on whether the condition is met. That depends on how the information was produced. An explicit causal model shows how a diagnostic can support the reactor conditional while leaving the probability of a restart unchanged. The same model also shows why that invariance does not determine the rest of the posterior. I introduce a domain Bayes factor that places materialization, Adams conditioning, and source-sensitive alternatives in a common framework. Trivalence's epistemological contribution is therefore not a uniquely privileged updating rule. It identifies a question that any adequate model of conditional learning must answer: does the evidence concern what happens under a condition, whether that condition obtains, or both?

1 Introduction

Suppose that a safety engineer tells us:

> If the reactor is restarted tonight, its core will overheat.

We trust the engineer and come to accept what she says. What, exactly, have we learned?

One natural answer is that we have learned what would happen on a restart. If the decision to restart will later be made by a fair coin toss, the engineer's diagnostic expertise need not tell us anything about the toss. Yet one prominent rule for learning conditionals says otherwise. The material counterpart of the engineer's statement is the proposition that

*Readers who would like to discuss the paper with the author may do so via this ChatGPT conversation.

either there will be no restart or there will be overheating. Conditioning on that proposition eliminates restart-without-overheating possibilities and renormalizes globally. From a uniform prior over the four combinations of restart and overheating, the probability of a restart falls from one-half to one-third.

This is not obviously irrational. An unsolicited warning can be evidence that its antecedent is a live possibility. A report may reveal the source's causal model, stable-property information, or reasons for mentioning one contingency rather than another. In such cases, learning the conditional may properly change the antecedent's probability. The mistake would be to suppose that the sentence alone decides whether it does.

Trivalent semantics brings the issue into focus. On a simple trivalent account, *if A, then B* is true when $A \wedge B$, false when $A \wedge \neg B$, and void when $\neg A$. The third value marks cases in which the issue posed by the conditional does not arise.[1] Its probability is the proportion of true cases among the cases in which it is not void. Thus confidence in $A \rightarrow B$ is $P(B \mid A)$, in accordance with Adams's Thesis.[2]

Recent work by Égré et al. (2025, 2026) combines this semantics with an elegant compositional probability theory. It also defines a rule of trivalent conditionalization. Under that rule, learning $A \rightarrow B$ has the same effect on unconditional probabilities as conditioning on $\neg A \vee B$. The result is a theorem of their proposed dynamics, not an algebraic error. Moreover, it follows from a general rule rather than merely from an informal inference about certainty.

The present paper asks a prior normative question: what makes that formal rule the right representation of *learning*? Égré, Rossi, and Sprenger note that the indicative and material conditionals have the same condition for certainty: both are certain when no probability remains on $A \wedge \neg B$. But a condition on admissible posteriors does not select one admissible posterior. Nor does assigning three truth values by itself specify how an information-producing process bears on those values. The semantics distinguishes success, failure, and non-application. It does not say whether receiving a report is evidence that the issue to which the report applies will arise.

My central claim is therefore one of *structured underdetermination*. Trivalent content supplies a distinguished partition of possibilities, and that partition reveals a dimension of evidence that bivalent materialization hides. But the semantics does not privilege a value on that dimension. An acquisition model must say whether the signal favors the conditional's domain, its void cases, or neither. A further assumption is needed to say how probability should move within each of the three cells. Conditional content, source structure, and

[1] The idea has a long history in work on conditional assertion and conditional bets; see Cooper (1968) and Belnap (1970). For recent probabilistic treatments, see Lassiter (2020) and Égré et al. (2026).

[2] See Adams (1975) and Edgington (1995).

conservativity do different work.

This conclusion changes the status of Domain Preservation. If receiving a signal is probabilistically independent of whether the conditional's domain is entered, Bayes's theorem requires the domain probability to be preserved. But that independence is relative to an agent's probabilistic model; it is not an intrinsic feature of a sentence or source. More importantly, preserving the domain does not determine the rest of the posterior. A natural engineer model will preserve the probability of a restart while changing beliefs about the reactor's disposition even at worlds where no restart occurs.

I make these points precise in three steps. First, I introduce a *domain Bayes factor* that measures how a learning signal bears on the conditional's domain. Materialization and Domain Preservation occupy different points in a continuous family. Second, I give an explicit causal model in which a diagnostic supports the reactor conditional while being domain-neutral for a principled reason. The model also exposes the limits of a simple domain-preserving update. Third, I explain what remains distinctively semantic: a compositional trivalent theory identifies the domain not only of simple conditionals but of arbitrary compounds. This yields invariance questions that cannot in general be read from a surface antecedent.

The resulting position is close in spirit to source-sensitive approaches in the existing literature. Douven (2012) argues that explanatory relations affect conditional learning, while Eva et al. (2020) distinguish interpreting evidence from updating on an interpretation and explicitly use an independent "disabler" to explain antecedent preservation. I do not claim to rediscover source sensitivity. The new claim is that even a unified, compositional trivalent semantics and probability theory leaves learning underdetermined in a systematic way. Trivalence identifies the missing coordinate, and thereby lets us diagnose what different update rules add to the semantics.

Section 2 reconstructs the materialization result and isolates the freedom left by certainty. Section 3 introduces domain evidence. Section 4 gives the causal model. Section 5 explains what trivalence contributes and distinguishes that contribution from earlier work. Section 6 treats minimum-change updating as a conservative completion rather than an autonomous theory of learning. A short representation result appears in the Appendix.

2 Certainty does not determine learning

Let W be a finite set of possible worlds, and let P be a probability distribution over them. A trivalent content E divides W into three cells:

$$E^{\mathsf{T}}, \qquad E^{\mathsf{F}}, \qquad E^{\mathsf{V}},$$

the worlds at which E is respectively true, false, and void. Write

$$D_E = E^{\mathsf{T}} \cup E^{\mathsf{F}}$$

for the *domain* of E: the worlds at which the issue settled by E arises. Its trivalent probability is

$$\pi_P(E) = \frac{P(E^{\mathsf{T}})}{P(E^{\mathsf{T}}) + P(E^{\mathsf{F}})}, \tag{1}$$

provided $P(D_E) > 0$. This is just ordinary probability measured within the nonvoid domain. If E is bivalent, there are no void worlds and (1) reduces to $P(E)$.

For a simple conditional with bivalent antecedent and consequent,

$$(A \to B)^{\mathsf{T}} = A \wedge B, \qquad (A \to B)^{\mathsf{F}} = A \wedge \neg B, \qquad (A \to B)^{\mathsf{V}} = \neg A.$$

Equation (1) then gives $\pi_P(A \to B) = P(B \mid A)$. The numerator measures the A-cases in which the conditional succeeds; the denominator measures all the cases in which it is tested. The $\neg A$-cases belong to neither.

The materialization result concerns what happens when an agent learns E. Égré et al. (2026, §6) define trivalent conditionalization by

$$\pi_P^E(H) = \pi_P(E \to H). \tag{2}$$

Their Proposition 5 establishes that formulas equivalent in the certainty-preserving logic induce the same update. The accompanying rationale is that learning E makes E certain; hence, if certainty in E and certainty in E' always coincide, learning either should yield the same posterior.

For a simple conditional, that principle yields materialization. Let

$$M_E = E^{\mathsf{T}} \cup E^{\mathsf{V}} = W \setminus E^{\mathsf{F}}$$

be E's *material shadow*. In the simple case, M_E is $\neg A \vee B$. Since $A \to B$ and $\neg A \vee B$ have the same falsifying worlds, they are equivalent in the certainty logic. Proposition 6 therefore implies that updating on the former has the same effect on the underlying distribution as ordinary conditioning on the latter:

$$P_{\mathsf{Mat}}^E(X) = P(X \mid M_E). \tag{3}$$

It removes the false worlds and preserves all prior probability ratios among the true and void

worlds that remain. Their Updating Theorem then shows that this result also minimizes every f-divergence from the prior subject to making E certain.[3]

The reactor case displays what this means. Let R say that the reactor is restarted and O that it overheats. Suppose initially that the four possibilities are equiprobable. Table 1 compares materialization with a domain-preserving, cell-conservative completion.

World	Prior	Material	Domain-preserving
$R \wedge O$	1/4	1/3	1/2
$R \wedge \neg O$	1/4	0	0
$\neg R \wedge O$	1/4	1/3	1/4
$\neg R \wedge \neg O$	1/4	1/3	1/4

Table 1: Two completions of the same certainty constraint.

Both updates make $P(O \mid R) = 1$. They differ in how they get there. Materialization eliminates $R \wedge \neg O$ and renormalizes globally, so $P(R)$ falls from 1/2 to 1/3. The domain-preserving update reallocates the probability of $R \wedge \neg O$ to $R \wedge O$ while leaving the no-restart worlds untouched, so $P(R)$ remains 1/2. Nothing in the demand for certainty chooses between these distributions.

The two updates therefore disagree neither about success nor about which propositions become certain. Both make the conditional and its material shadow certain. They disagree about how probability should be distributed among possibilities compatible with those certainties. That is a difference in evidence even when it is not a difference in categorical commitment.

The point is general. To be certain of a trivalent content E is only to assign no probability to E^{F}. That constraint does not settle the division of probability between E^{T} and E^{V}. Yet the distinction between those cells is exactly what separates an indicative conditional from its material shadow. Trivalent conditionalization makes true and void worlds dynamically alike in one respect: after falsity is eliminated, their prior ratio is preserved. That is a substantive bridge from content to change, not a consequence of certainty alone.

This observation does not refute trivalent conditionalization as a formal proposal. Égré, Rossi, and Sprenger define the rule independently, and may defend it as the correct content-level norm. The narrower conclusion is that neither certainty equivalence nor the trivalent semantics uniquely supports that norm. To see the available alternatives, we need to represent how a signal bears on the domain.

[3]These are Definition 8, Propositions 5 and 6, and Theorem 4 of Égré et al. (2026). The Reduction Theorem extends the result to arbitrary compounds by identifying a simple conditional with the same trivalent semantic profile.

The difficulty is specific to the attempted extension of Bayesian learning. For a bivalent proposition, the worlds at which the proposition is true are also all the worlds at which it is not false. This supplies the familiar idealization of learning the proposition: condition on its truth set. A trivalent content separates those roles. Conditioning on E^{T} excludes both falsity and voidness; conditioning on $E^{\mathsf{T}} \cup E^{\mathsf{V}}$ yields materialization; conditioning on an actual signal can produce either of these effects or neither. The semantic object therefore does not identify a unique evidence event.

Call this the *trivalent bridge problem*. A theory may solve it by adding a norm that maps trivalent content to an update, as trivalent conditionalization does. Or it may model the event through which the content is acquired. But the bridge is an additional part of the epistemology. In the bivalent case its role is easy to miss because truth and non-falsity coincide; trivalence makes the choice visible.

3 The domain-evidence coordinate

Let

$$t = P(E^{\mathsf{T}}), \qquad n = P(E^{\mathsf{F}}), \qquad v = P(E^{\mathsf{V}}), \qquad r = t + n = P(D_E),$$

and assume for now that all three cells have positive probability. Any posterior Q with $Q(D_E) > 0$ can be described in part by two coordinates:

$$s = \pi_Q(E), \qquad r' = Q(D_E).$$

The first says how strongly the posterior favors truth over falsity within the domain. The second says how strongly it favors entering the domain rather than remaining in the void cases. Together they fix the posterior cell masses:

$$Q(E^{\mathsf{T}}) = sr', \qquad Q(E^{\mathsf{F}}) = (1-s)r', \qquad Q(E^{\mathsf{V}}) = 1 - r'. \tag{4}$$

They do not fix how probability is distributed inside the cells.

Now let S be the event that an agent receives a report, diagnostic output, or other learning signal. Define the signal's *domain Bayes factor* by

$$\kappa_S = \frac{P(S \mid D_E)}{P(S \mid E^{\mathsf{V}})}. \tag{5}$$

when the denominator is positive. Zero likelihoods yield the corresponding limiting values.

Bayes's theorem gives

$$P(D_E \mid S) = \frac{\kappa_S r}{\kappa_S r + v}. \tag{6}$$

Thus $\kappa_S > 1$ means that the signal supports entering the domain, $\kappa_S < 1$ means that it supports voidness, and $\kappa_S = 1$ means that it is neutral between them.

Domain Preservation is the neutral case:

$$\kappa_S = 1 \quad \Longrightarrow \quad P(D_E \mid S) = P(D_E). \tag{7}$$

For a simple $A \to B$, this is antecedent preservation. Such preservation is familiar from Adams conditioning and earlier preservation accounts.[4] Equation (7) does not yet give an updating rule. It fixes one marginal of the posterior.

Materialization occupies a different point on the same scale. A formal likelihood representation of categorical materialization assigns equal positive likelihood to E^{T} and E^{V} and zero likelihood to E^{F}. Consequently,

$$P(S_{\mathsf{Mat}} \mid D_E) = \frac{t}{r}, \qquad P(S_{\mathsf{Mat}} \mid E^{\mathsf{V}}) = 1,$$

and hence

$$\kappa_{S_{\mathsf{Mat}}} = \frac{t}{r} = \pi_P(E). \tag{8}$$

Unless E was already certain, materialization has a domain Bayes factor below one. This is a description of the rule's probabilistic effect. It is not the claim that proponents of materialization are committed to a psychological source with these literal reporting propensities. They may instead offer the rule as a normative bridge governing content-level learning.

The comparison reveals a continuum rather than a forced choice. The following elementary result records the extent of the freedom.

Theorem 1 (Cell-coarse realization)**.** *Suppose $t, n, v > 0$. For every target $s \in [0, 1]$ and every finite $\kappa > 0$, there is a signal whose likelihood is constant within each semantic cell, whose domain Bayes factor is κ, and after which the trivalent probability of E is s.*

Proof. Set $r_\kappa = \kappa r/(\kappa r + v)$ and assign the three cells relative likelihoods

$$\lambda_{\mathsf{T}} = \frac{s r_\kappa}{t}, \qquad \lambda_{\mathsf{F}} = \frac{(1-s) r_\kappa}{n}, \qquad \lambda_{\mathsf{V}} = \frac{1 - r_\kappa}{v}.$$

Their prior-weighted sum is one, so Bayes's theorem gives posterior cell masses $s r_\kappa$, $(1-s) r_\kappa$, and $1 - r_\kappa$. Multiplying all three numbers by a sufficiently small common constant turns

[4] See Adams (1975), Bradley (2000), and Douven (2012). The present terminology emphasizes that, for complex trivalent contents, the preserved event need not be a surface antecedent.

them into probabilities of a signal without changing the posterior or any likelihood ratio. Equations (5) and (6) then give the required domain Bayes factor. □

The theorem is a possibility result, not a vindication of any such channel. Likelihoods can be reverse-engineered for any absolutely continuous posterior. Its point is negative: neither trivalent content nor the demand for success selects one member of the family. Moreover, signals can distinguish worlds within a cell, in which case the same s and κ are compatible with many further posteriors. The domain effect and within-cell change are both left open.

There is an important qualification. Source neutrality is prior-relative. $P(S \mid D_E)$ averages the signal's likelihood over the true and false cells using the agent's conditional prior. In the categorical, cell-coarse case, $P(S \mid E^{\mathsf{F}}) = 0$, and neutrality requires

$$\frac{P(S \mid E^{\mathsf{T}})}{P(S \mid E^{\mathsf{V}})} = \frac{1}{P(E^{\mathsf{T}} \mid D_E)}. \tag{9}$$

A single channel with fixed cellwise likelihoods will not satisfy this equation for arbitrary priors. It is therefore misleading to call neutrality an intrinsic property of a source. It is an independence relation between a signal and a domain in a probabilistic model.

This does not make neutrality incapable of independent support. Experimental randomization or causal structure can entail the relevant independence for a family of priors that respect that structure. But the structure must be stated. The next section supplies it for the reactor case.

4 A generative model of conditional evidence

The four actual-value combinations in Table 1 are enough to display two posteriors, but not enough to explain how an engineer could know what a restart would produce. For that we need a variable representing the reactor's mechanism.

Let R say that the reactor is restarted, M that its present mechanism would produce overheating on a restart, O that overheating actually occurs, and S that the engineer's diagnostic returns a positive result. Suppose that the restart is fixed by a later randomizer, independently of M; the diagnostic depends on M but not on the randomizer; and overheating occurs exactly when the reactor is restarted with the overheating mechanism. The causal structure is

$$R \longrightarrow O \longleftarrow M \longrightarrow S,$$

with joint distribution

$$P(R, M, O, S) = P(R)P(M)P(S \mid M)\mathbf{1}[O = R \wedge M]. \tag{10}$$

Assume the diagnostic is informative: $P(S \mid M) > P(S \mid \neg M)$, with $0 < P(M) < 1$.

This model gives the desired result for a principled reason. Because R is randomized independently of M and the diagnostic depends only on M,

$$P(R \mid S) = P(R). \tag{11}$$

At the same time,

$$P(O \mid R, S) = P(M \mid S) > P(M) = P(O \mid R). \tag{12}$$

The diagnostic therefore raises the probability of the conditional within its restart domain while leaving the probability of entering that domain fixed. If the diagnostic is perfectly reliable, $P(M \mid S) = 1$, and the conditional $R \to O$ becomes certain.

The same model exposes a limitation of the simple update in Table 1. At no-restart worlds the conditional is void, but the diagnostic still informs us about the mechanism:

$$P(M \mid \neg R, S) = P(M \mid S) > P(M).$$

Thus the signal preserves the total probability of the void cell while changing probability ratios *within* it. A total update that leaves every void world untouched would discard information the source actually provides. Domain neutrality licenses Domain Preservation, not cellwise invariance.

This is structurally close to the analysis of the Sundowners case by Eva et al. (2020). Their "disabler" variable records an independent reason why rain would prevent sundowners; updating on that reason together with the material conditional preserves the probability of rain. The reactor mechanism M plays the same role. That earlier analysis anticipates the appeal to source independence, and the present proposal should not be advertised as a new justification of antecedent preservation. Its role here is different: it shows what an adequate acquisition model adds to a trivalent semantic partition, and why a preserved domain still leaves the full posterior underdetermined.

Edgington's coin illustrates the same point. If whether the coin is flipped is randomized independently of its being double-headed or double-tailed, and an informant reports after inspecting the coin's type, the report is neutral about the flip. It nevertheless distinguishes the two no-flip possibilities inside the void cell and can make the coin's type certain. The materialization value of 2/3 and the stable-type value of 1 correspond to different models of

the information, not different computations from a fully specified learning event.[5]

Other causal structures produce other domain Bayes factors. A speaker may be more likely to issue a warning when she expects its antecedent to occur. A common cause may influence both the source's evidence and domain entry. Learning the conditional may also alter an agent's later decision about whether to make the antecedent true. In these cases S and D_E are not independent, and Domain Preservation is not warranted.[6]

5 What trivalence contributes

The argument so far denies that semantics selects a privileged acquisition model. It does not make semantics idle. Trivalence determines the partition relative to which different dimensions of learning can be separated. The division of labor is summarized in Table 2.

Question	Formal object	Supplied by
Where does the content succeed, fail, or not apply?	$E^{\mathsf{T}}, E^{\mathsf{F}}, E^{\mathsf{V}}$	semantics
Does the signal favor the content's domain or its void cases?	domain Bayes factor κ_S	acquisition model
Does the signal carry finer information within a cell?	statewise likelihoods	acquisition model or a conservativity assumption

Table 2: Three questions in conditional learning.

For simple $A \to B$, the domain is A, and the domain-evidence question is the familiar question whether the evidence changes the probability of the antecedent. The advantage of a compositional trivalent semantics appears with complex contents. The Reduction Theorem of Égré et al. (2026) associates any compound E with a simple conditional $A_E \to B_E$ having the same trivalent profile. Its domain is the bivalent event A_E. The theory therefore tells us which event plays the role of applicability even when there is no single surface antecedent.

Consider, for example,

$$E = (A \to B) \wedge (C \to D).$$

Under the Cooper conjunction used by Égré, Rossi, and Sprenger, this compound is void only when both component antecedents are false. Its domain is therefore $A \vee C$. A domain-neutral acquisition model for the conjunction preserves $P(A \vee C)$, not necessarily $P(A)$ and $P(C)$ separately. A diagnostic might change the relative probability of A-only and C-only cases

[5] See Edgington (1995) and Égré et al. (2026).

[6] For source and context effects, see Collins et al. (2020); for causal approaches to conditional learning, see Vandenburgh (2021).

while remaining neutral about whether at least one test occurs. This invariance cannot be read off from Adams conditioning applied independently to the two surface conditionals.

The consequence is substantive even though it is conditional on the semantics. Different trivalent treatments of conjunction or nesting may assign different domains to the same compound. They thereby identify different independence questions for acquisition models and different marginals as candidates for preservation. Dynamic data cannot by themselves choose a semantics, but the semantic theory determines what counts as domain evidence.

This locates the contribution relative to earlier work. Adams conditioning and Bradley (2000) supply preservation principles for simple conditionals. Douven (2012) shows that explanatory considerations can move antecedent probabilities. Eva et al. (2020) provide an explicit two-stage picture: interpret the total evidence, including causal and pragmatic structure, and then update on that interpretation. The present account accepts those lessons. It adds three claims. First, a fully compositional trivalent semantics still does not determine an update rule. Second, it supplies a canonical domain coordinate for arbitrary compounds, not merely simple antecedents. Third, the domain Bayes factor shows exactly what materialization, preservation, and other rules say about that coordinate.

The last point also clarifies the disagreement with Égré, Rossi, and Sprenger. Their trivalent conditionalization is a coherent normative bridge from semantic content to posterior belief. Proposition 6 and the Updating Theorem establish important consequences of adopting it. What they do not establish is that the bridge is forced by the truth conditions, by certainty, or by minimum change without a prior specification of the evidential constraints. A defender of materialization can accept this and offer an independent content-level norm. A source-sensitive theorist can instead let the acquisition model determine κ_S and any within-cell information.

Trivalence's epistemological payoff is thus diagnostic rather than dictatorial. It shows why learning what happens under a condition and learning whether the condition obtains are distinct dimensions. It does not say that the second dimension must remain fixed.

6 Conservative completions and limits

Sometimes an agent or modeler specifies only the two coordinates s and r' and has no further information within the semantic cells. A natural conservative completion preserves prior probability ratios within each cell. Assume first that $t, n, v > 0$; if a semantic cell is empty,

its line is simply omitted:

$$P_{\mathsf{C}}^{E;s,r'}(w) = \begin{cases} P(w)\dfrac{sr'}{t}, & w \in E^{\mathsf{T}}, \\ P(w)\dfrac{(1-s)r'}{n}, & w \in E^{\mathsf{F}}, \\ P(w)\dfrac{1-r'}{v}, & w \in E^{\mathsf{V}}. \end{cases} \tag{13}$$

This is the update displayed in Table 1 when $s = 1$ and $r' = r$. It should not be confused with the full update generated by the causal model in Section 4. Equation (13) represents evidence that is coarse-grained exactly to the three semantic cells, or a modeler's decision to add no changes beyond the specified cell masses.

The representation theorem in the Appendix gives this modest claim a familiar form. Once s and r' have fixed all three cell totals, (13) uniquely minimizes every strictly convex f-divergence in the class defined there from the prior. This is an application of Jensen's inequality. It does not independently justify either target. Minimum change completes an acquisition model; it does not replace one.[7]

Several familiar rules are special cases. If $r' = r$, the completion preserves the domain. For a simple conditional it is Adams conditioning: it changes $P(B \mid A)$ to s, preserves $P(A)$, and preserves prior ratios within the three cells. If E is bivalent and $s = 1$, the void line is omitted and the completion reduces to ordinary Bayesian conditionalization. Categorical materialization is obtained by setting

$$s = 1, \qquad r' = \frac{t}{t+v}.$$

These rules differ in their substantive targets, not in the mathematical principle used to fill out the posterior.

The framework has ordinary support limitations. Equation (1) is undefined when $Q(D_E) = 0$. Thus the claim that certainty is equivalent to assigning zero probability to E^{F} requires either the additional condition $Q(D_E) > 0$ or a stipulated value for the all-void case. Égré, Rossi, and Sprenger adopt such a stipulation; it should not be allowed to pass unnoticed. When the prior assigns zero probability to a cell that learning must revive, no update absolutely continuous with respect to the prior will suffice. Popper functions, lexicographic probabilities, imaging, or prior expansion may be needed.[8]

Nor is a true/false/void partition a complete representation of every conditional-learning

[7]The same moral is familiar from the Judy Benjamin literature; see van Fraassen (1981); Douven and Romeijn (2011).

[8]See Günther (2018) and Günther and Trpin (2023).

problem. Causal and imaging approaches encode relations among worlds that a three-cell partition omits. Dynamic accounts can represent conditional probability through hypothetical update rather than the probability of a trivalent content (Ciardelli and Ommundsen, 2024). The present argument is internal to the trivalent project: if conditionals are represented by three-valued contents, their semantics structures but does not settle their dynamics.

7 Conclusion

What one learns from a conditional depends on more than the conditional's truth conditions. A trivalent content distinguishes the possibilities in which it succeeds, fails, and does not apply. Certainty constrains the false cell but leaves open the balance between the domain and the void cases. The same conditional can therefore be learned through signals that raise, lower, or preserve the probability of its domain.

The domain Bayes factor makes that freedom explicit. Materialization assigns a particular domain effect; Domain Preservation assigns another. Neither follows from trivalence alone. An explicit reactor model shows how randomized domain entry and a diagnostic of an independent mechanism can justify preservation. It also shows that preservation fixes only a marginal: the diagnostic may carry substantial information within the void cell.

This revision of the claim leaves a genuine role for semantics. A compositional trivalent theory determines the applicability domain of complex contents and hence identifies the event relative to which source effects must be assessed. For compounds, that event need not be any surface antecedent. Semantics tells us what the domain question is; an acquisition model tells us how the evidence answers it.

Trivalent conditionalization remains a coherent proposal, and its materialization theorem remains formally significant. But a semantic theory does not become a theory of learning merely by adding a rule called conditionalization. The rule needs a bridge from content to evidence. For conditionals, making that bridge explicit is not an optional complication. It is part of saying what was learned.

A Representation theorem

Let $f : [0, \infty) \to \mathbb{R}$ be continuous and strictly convex, with $f(1) = 0$, and define

$$D_f(Q, P) = \sum_{w \in W} P(w) f\left(\frac{Q(w)}{P(w)}\right). \tag{14}$$

Kullback–Leibler divergence is the special case given by the continuous extension of $f(x) = x \log x$ with $f(0) = 0$. No differentiability at zero is assumed.

Theorem 2 (Representation)**.** *Let W be finite, let P have full support, and suppose the three semantic cells have positive prior probability. For any $s \in [0,1]$ and $r' \in (0,1)$, $P_{\mathsf{C}}^{E;s,r'}$ is the unique distribution that minimizes every strictly convex divergence $D_f(Q,P)$ subject to*

$$\pi_Q(E) = s, \tag{15}$$

$$Q(D_E) = r'. \tag{16}$$

Proof of Theorem 2. The two constraints fix the total posterior mass of all three cells. Since $Q(D_E) = r'$ and $\pi_Q(E) = s$,

$$Q(E^{\mathsf{T}}) = sr', \qquad Q(E^{\mathsf{F}}) = (1-s)r', \qquad Q(E^{\mathsf{V}}) = 1 - r'.$$

It remains to determine how each fixed mass is distributed within its cell. Consider any cell C with $P(C) > 0$ and required posterior mass $Q(C) = m$. If $m = 0$, every admissible Q vanishes throughout C. Suppose, then, that $m > 0$. Its contribution to the divergence is

$$P(C) \sum_{w \in C} P(w \mid C) f\left(\frac{m}{P(C)} \frac{Q(w \mid C)}{P(w \mid C)}\right).$$

The $P(\cdot \mid C)$-weighted mean of $Q(w \mid C)/P(w \mid C)$ is one. By strict convexity and Jensen's inequality, the displayed expression is uniquely minimized when

$$\frac{Q(w \mid C)}{P(w \mid C)} = 1$$

for every $w \in C$. Equivalently, $Q(w) = P(w)m/P(C)$ throughout C. Applying this result to the true, false, and void cells yields exactly (13). At a boundary value of s, the zero-mass observation applies to one of the first two cells and the Jensen argument applies to the remaining cells. □

References

Adams, Ernest W. 1975. *The Logic of Conditionals.* Dordrecht: Reidel.

Belnap, Nuel D. 1970. Conditional assertion and restricted quantification. *Noûs* 4(1): 1–12.

Bradley, Richard. 2000. A preservation condition for conditionals. *Analysis* 60(3): 219–222.

Bradley, Richard. 2005. Radical probabilism and Bayesian conditioning. *Philosophy of Science* 72(2): 342–364.

Ciardelli, Ivano, and Adrian Ommundsen. 2024. Probabilities of conditionals: Updating Adams. *Noûs* 58(1): 26–53.

Collins, Peter J., Karolina Krzyżanowska, Stephan Hartmann, Gregory Wheeler, and Ulrike Hahn. 2020. Conditionals and testimony. *Cognitive Psychology* 122: 101329.

Cooper, William S. 1968. The propositional logic of ordinary discourse. *Inquiry* 11(1–4): 295–320.

Douven, Igor. 2012. Learning conditional information. *Mind & Language* 27(3): 239–263.

Douven, Igor, and Jan-Willem Romeijn. 2011. A new resolution of the Judy Benjamin problem. *Mind* 120(479): 637–670.

Edgington, Dorothy. 1995. On conditionals. *Mind* 104(414): 235–329.

Égré, Paul, Lorenzo Rossi, and Jan Sprenger. 2025. Certain and uncertain inference with indicative conditionals. *Australasian Journal of Philosophy* 103(3): 569–596.

Égré, Paul, Lorenzo Rossi, and Jan Sprenger. 2026. Probability for trivalent conditionals. *Mind*, in press.

Eva, Benjamin, Stephan Hartmann, and Soroush Rafiee Rad. 2020. Learning from conditionals. *Mind* 129(514): 461–508.

Günther, Mario. 2018. Learning conditional information by Jeffrey imaging on Stalnaker conditionals. *Journal of Philosophical Logic* 47(5): 851–876.

Günther, Mario, and Borut Trpin. 2023. Bayesians still don't learn from conditionals. *Acta Analytica* 38(3): 439–451.

Lassiter, Daniel. 2020. What we can learn from how trivalent conditionals avoid triviality. *Inquiry* 63(9–10): 1087–1114.

van Fraassen, Bas C. 1981. A problem for relative information minimizers in probability kinematics. *British Journal for the Philosophy of Science* 32(4): 375–379.

Vandenburgh, Jonathan. 2021. Conditional learning through causal models. *Synthese* 199: 2415–2437.