

The Wormwood of Penal Substitution and the Miscarriage of Justice — 34

Phil Stilwell

Independent Scholar

ORCID: [0009-0009-3725-8682](https://orcid.org/0009-0009-3725-8682)

Revised August 2026

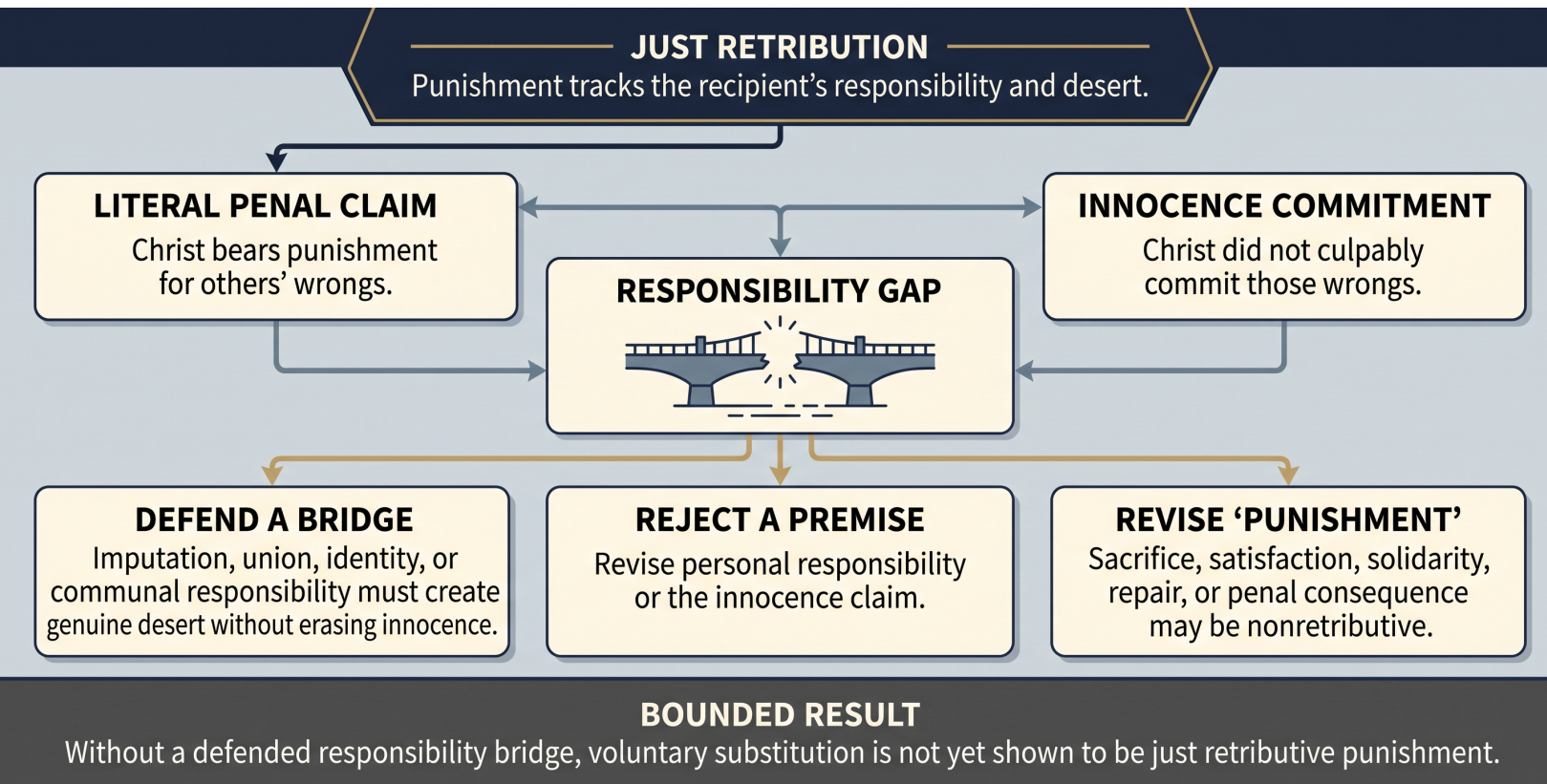
Central claim

Literal retributive penal substitution says that Christ, though innocent of others' wrongdoing, justly receives the punishment they deserve. Just retribution, however, tracks the recipient's culpable responsibility and desert; consent can authorize suffering but cannot supply either. The theory therefore faces a responsibility-bridge dilemma: it must show how imputation, union, identity, or communal responsibility makes Christ genuinely deserving without erasing innocence, reject a premise, or revise "punishment" into sacrifice, satisfaction, solidarity, repair, or penal consequence. Without that bridge, voluntary substitution is not shown to be just retributive punishment; Christian atonement and nonretributive substitution remain open.

Preview:

Literal claim	Responsibility rule	Required response	Bounded result
Christ bears punishment for others' wrongs while remaining innocent.	Just retribution tracks the recipient's culpable responsibility and desert.	Defend a bridge, reject a premise, or revise what "punishment" means.	Without the bridge, voluntary substitution is not shown to be just retribution; other atonement models remain open.

Visual preview:



The diagram isolates the paper's governing dilemma: just retribution tracks recipient responsibility; the literal penal claim plus the innocence commitment create a responsibility gap. A defender must supply a bridge, reject a premise, or revise punishment. The diagram does not rule out Christian atonement or nonretributive substitution.

Abstract

Literal retributive penal substitution combines three claims: Christ is justly punished for others' wrongdoing; just retribution tracks the recipient's culpable responsibility and desert; and Christ is innocent. Together they create a responsibility-bridge dilemma. A defender must explain how Christ acquires the responsibility and desert that retribution requires without erasing his innocence, reject one of those commitments, or revise "punishment" into a nonretributive category. Consent authorizes sacrifice but does not create desert; civil analogies transfer costs more readily than penal censure; and imputation, union, enlarged identity, vicarious punishment, and communal responsibility each need a distinct responsibility-making relation. Scriptural plurality, historical development, scapegoat analysis, and qualitative evidential comparison raise or clarify the burden without independently deciding it. The bounded result is that, without a defended bridge, voluntary substitution is not shown to constitute just retributive punishment. This conclusion does not disprove Christian theism, the religious significance of the cross, substitution in every sense, or nonretributive theories of atonement.

Keywords: penal substitution; atonement; punishment; personal responsibility; desert; imputation; justice; scapegoat

1. The Exact Claim Under Pressure

The broad language of substitution covers several importantly different ideas. One person can pay another's debt, absorb another's loss, repair damage another caused, represent a community, rescue someone at great personal cost, or suffer because of another's wrongdoing. None of those descriptions alone entails that the substitute is *punished*. The morally difficult claim is more specific: Christ receives the retributive punishment deserved by human wrongdoers, and this transfer satisfies perfect justice.

That claim has prominent defenders, serious philosophical refinements, and targeted critics. Classic Protestant presentations describe Christ as bearing the penalty due to sinners, while contemporary work distinguishes punishment from the mere infliction of pain, stresses Christ's voluntary agency, and develops imputation or union as a response to the innocence objection (Craig, 2018a, 2018b; Kyle, 2013; Stott, 1986; Woznicki, 2019). The present argument therefore does not attack the slogan "Christ died for us" or assume that every penal-substitution theory is identical. It targets the version with three commitments:

1. Christ undergoes punishment, not merely costly rescue, suffering, or symbolic identification.
2. The punishment is retributive: it answers wrongdoing because punishment is deserved.
3. Christ is innocent of the human wrongdoing for which he is punished.

A *God-concept* is a specified conception of a deity, including the moral attributes and redemptive actions assigned to that deity. The Christian God-concept examined here includes perfect justice and the innocence of Christ. *Theism* is the broader claim that at least one deity exists. An internal inconsistency in one atonement explanation would not refute theism, and it would not show that every Christian understanding of the cross fails.

The governing question is simple: what makes punishment of this particular person for that wrongdoing just? "Substitution" describes the proposed transfer. "Atonement" names its intended reconciling function. "Perfect justice" intensifies the need for a justification. None of those labels supplies the missing relation between the punished person and the wrong.

Call this the *responsibility-bridge dilemma*. Every major route considered below faces the same two questions. Does it supply a relation strong enough to make Christ genuinely deserving of retribution while preserving his innocence? Or does it coherently revise the event as sacrifice, satisfaction, solidarity, repair, or consequence rather than retributive punishment? The paper's supporting historical, scriptural, scapegoat, and evidential strands matter insofar as they clarify or alter that single burden.

The responsibility argument

1. Just retributive punishment of a person for a wrong requires that person to deserve punishment for that wrong through culpable responsibility.
2. Christ did not culpably commit or participate in human wrongdoing; absent a further responsibility-making relation, he does not deserve retribution for it.
3. Literal retributive penal substitution says that Christ is justly punished for human wrongdoing.
4. These three claims cannot all be true. A defender must reject the responsibility condition, reject Christ's innocence, or add a bridge by which Christ acquires genuine responsibility and desert; redescribing the event as nonretributive suffering changes the target claim.

The argument is not that substitution is logically impossible under every imaginable metaphysics. It identifies a burden. An account of imputation, union, corporate identity, or voluntary representation could in principle supply a bridge. But the bridge must do more than repeat that guilt is “counted” to Christ. It must explain why the counting makes retributive treatment deserved rather than merely authorized, useful, loving, or accepted.

2. The Governing Standard: Punishment, Desert, and Responsibility

2.1 Why hard treatment is not yet punishment

Punishment is not simply suffering caused by an authority. Surgery, quarantine, rescue work, restitution, athletic training, and self-sacrifice can all impose severe burdens without punishing the person burdened. Philosophical accounts differ over punishment's ultimate justification, but retributive and communicative theories characteristically connect it to responsible wrongdoing. Hart's responsibility-sensitive analysis, Feinberg's account of punishment as an expression of condemnation, Duff's communicative account, and Moore's desert-centered retributivism all make the punished person's answerability morally central (Duff, 2001; Feinberg, 1970; Hart, 1968; Moore, 1997).

The point does not require treating modern criminal law as a perfect mirror of divine justice. Human institutions are fallible, politically constituted, and constrained by practical rules. The comparison isolates a moral grammar: censure addresses an agent as the doer of a wrong, and retribution gives someone the hard treatment that person's culpable conduct deserves. When that agent-relative link disappears, the treatment may serve deterrence, compensation, ritual, expression, solidarity, or rescue. Calling it retribution does not recreate the missing desert.

This distinction also prevents an easy equivocation. Suppose Christ's suffering vividly condemns human sin. That may give the event an expressive function. But condemning *the wrong* through an innocent person's suffering is not yet punishing *the innocent person for committing it*. If the censure is not directed at Christ as a responsible wrongdoer, one element of punishment has been withdrawn. If it is directed at him, innocence makes the direction morally suspect.

2.2 Personal responsibility is a benchmark, not a claim of perfect uniformity

Legal and moral practices contain collective, corporate, and vicarious forms of liability. Parents sometimes bear costs generated by children; employers may answer for employees; corporations may be sanctioned for actions distributed across many people. These practices show that responsibility is relational and institutionally mediated. They do not establish that penal desert can be assigned to any willing substitute.

Even vicarious liability normally depends on a role, authorization, benefit, control, participation, or constitutive relationship. It is controversial precisely when sanctions detach from those responsibility-making features.

Corporate punishment also presupposes an argument that the organization is an agent or bearer of responsibility, not simply a convenient innocent target. Cawdron’s recent communal account is valuable here because it tries to articulate a form of group non-agential moral responsibility rather than assuming that proximity alone transfers guilt (Cawdron, 2024). Its care confirms the point: a relation that bears moral weight must be shown.

Personal responsibility should therefore be used as a strong public benchmark, not as an alleged universal description of every legal system. Exceptions matter. Yet an exception is not self-justifying. A penal-substitution theory must identify what makes the relation between Christ and wrongdoers strong enough to ground punishment, and why that relation transfers desert without converting innocence into culpability by stipulation.

Burden	What it can show	What it cannot show by itself	Argumentative force
Transferred desert	Retribution normally follows the culpably responsible person	That no enlarged identity or communal relation could ever ground responsibility	Central
Voluntary suffering	Christ accepts rather than merely endures the burden	That consent creates culpability for another’s act	Central
Scapegoat comparison	Innocent suffering can stabilize a community while misplacing blame	That every reading of the cross repeats scapegoating	Supporting
Civil/criminal contrast	Costs and debts transfer more readily than penal condemnation	That human legal rules directly bind divine action	Supporting
Scriptural plurality	Christian texts use multiple atonement images	That penal readings are textually impossible or contradictory	Contextual
History and comparison	Penal formulations are one developed way of organizing the material	That later development is false or that cultures share one uniform norm	Contextual
Evidential comparison	A theory loses explanatory force when it requires an unargued moral exception	A numerical posterior probability	Contextual

The table ranks the paper’s strands by their role in the responsibility-bridge dilemma. The responsibility and consent arguments establish the governing test. The remaining considerations ask whether a proposed route supplies the needed relation or instead supports a nonretributive account. They are not separate proofs.

3. Consent Authorizes Sacrifice; It Does Not Create Desert

Voluntariness answers an important objection. If Christ freely accepts suffering, the event differs morally from a ruler seizing an uninvolved victim. Love can make self-sacrifice admirable, and consent can remove a wrong that would consist in coercing the sufferer. Penal-substitution defenders are therefore right to insist that a purely involuntary-victim picture is inadequate.

But consent changes the wrong of imposition; it does not change who committed the original wrong. A volunteer may pay a fine, compensate a victim, donate an organ, enter danger, or bear social costs. None of these acts makes the volunteer culpable for what another person did. If a judge imprisoned an innocent volunteer in place of a convicted offender, willingness would not make the sentence retributively deserved. It might make the volunteer heroic. It would not make the volunteer guilty.

This is the distinction between *permission* and *desert*. Consent can give another party permission to impose a burden that would otherwise violate the consenter’s rights. Retributive desert says something further: the

suffering is a fitting response to the recipient's culpable wrongdoing. The first is controlled by the sufferer's will; the second is controlled by the moral relation between agent and act.

The distinction creates a dilemma for claims that voluntariness makes the punishment just. If Christ's consent is sufficient because the suffering is an accepted sacrifice, then the explanation is sacrificial rather than retributive. If retributive justice requires the suffering independently of consent, the account still needs desert. Combining consent with the assertion of punishment does not bridge the categories.

This objection is not defeated by the thought that the specified Christian deity and Christ share one divine will. That may deepen the voluntariness of the redemptive act and block a picture of one divine person coercing another. It still does not show that the human wrongdoing for which punishment is imposed is culpably Christ's. Trinitarian unity may answer *who authorizes the sacrifice*; it does not, without an additional account, answer *who deserves retribution for the wrong*. Consent therefore addresses one side of the responsibility-bridge dilemma while leaving the desert side unresolved.

4. The Scapegoat Warning—and Its Limit

The scapegoat comparison has rhetorical power because communities often secure unity by placing disorder, pollution, or guilt on an expendable victim. Penal substitution can appear to sanctify that pattern: wrongdoing is resolved when an innocent sufferer receives the violence that the guilty avoid. The comparison is especially sharp when suffering itself is said to be required before mercy can become just.

Yet the analogy must be disciplined. The live goat in Leviticus 16 symbolically bears Israel's iniquities into the wilderness; it is not a modern criminal defendant. Christian readings of the cross also differ. Heim, drawing on Girardian themes, construes the cross as exposing and overcoming the scapegoat mechanism rather than as divine endorsement of it (Heim, 2006). On that view, the victim's vindication reveals the community's violence. A critic cannot simply assume that every atonement account repeats the mechanism it may instead be diagnosing.

The analogy should therefore function as a warning test:

1. Does the account treat an innocent person's suffering as intrinsically required before forgiveness is permissible?
2. Is blame genuinely located in the punished person, or is condemnation theatrically displaced onto that person?
3. Does resurrection and vindication reverse the community's judgment, or does the account still call the victim's penal suffering retributively deserved?

If an account says that Christ enters human violence, exposes it, repairs its damage, and reconciles enemies, the scapegoat objection may miss its target. If the account says perfect justice requires that an innocent person receive the guilty persons' retribution, the warning lands. The comparison therefore returns to the same test: it either identifies a responsibility-making relation or supports a nonretributive reading; resemblance alone supplies no bridge.

5. Civil Analogies Transfer Costs, Not Penal Responsibility

Substitution is ordinary in civil and economic life. A guarantor pays a debt; an insurer covers a loss; a benefactor pays another's fine; a representative acts for a group. These examples make third-party satisfaction intelligible. They show that obligations, costs, and repairs can be discharged by someone other than the original wrongdoer.

They do not by themselves justify transferred punishment. Debt concerns what is owed; punishment concerns who is answerable. If a friend pays the amount I owe, the creditor can be made whole without pretending that the friend incurred my debt through misconduct. If a friend serves my prison sentence, the state has not merely received an equivalent quantity of hardship. It has condemned the wrong person. The personal character of penal censure is precisely what prevents punishment from functioning like a fungible payment.

Lewis pressed a related concern by asking whether common Christian belief really matches literal penal substitution once the elements of punishment are made explicit (Lewis, 1997). Murphy proposes “vicarious punishment” as a more careful category, and Craig answers coherence objections by distinguishing the object of condemnation and by defending legal imputation (Craig, 2018b; Murphy, 2009). These exchanges improve the debate because they reveal that the label is not enough. One must specify whether Christ is condemned, whether the suffering counts as punishment without condemnation of Christ, and what makes any legal imputation morally truth-tracking rather than fictional.

Legal fiction can coordinate institutions for good practical reasons. It does not automatically generate the underlying fact it treats as present. Declaring a corporation a “person” may allocate rights and duties; declaring an innocent individual “guilty” would not make punishment deserved. If divine imputation is not a fiction but a revelation of a deeper union, the theory must explain that union. If it is a purposive legal treatment despite the absence of culpability, the defense has shifted from desert to authorization or beneficial outcome.

6. The Reply Test: Does It Supply Responsibility or Revise Punishment?

The responsibility argument is useful only if it survives serious replies. Four routes deserve special attention. Each is more sophisticated than the caricature of an angry ruler punishing a random victim. Each therefore receives the same two-question test: does it supply a responsibility-making relation strong enough for retributive desert while preserving innocence, or does it revise the event into a coherent nonretributive form? Their answers and costs differ.

6.1 Legal imputation and vicarious punishment

On an imputation account, human guilt is legally reckoned to Christ, who voluntarily bears the sanction. Craig argues that punishment need not express the judgment that Christ is personally immoral; it can express condemnation of sin while Christ serves as a legally liable substitute (Craig, 2018a, 2018b). This blocks the claim that the theory must attribute a vicious character or wrongful action to Christ.

The remaining question is why reckoning is just. A legal status can be assigned, but retributive desert purports to track moral reality rather than administrative convenience. If imputation transfers liability without culpability, the suffering may be authorized vicarious punishment but not deserved because of Christ’s wrongdoing. If it transfers culpability, the innocence commitment is threatened. The theory needs an intermediate category—genuine responsibility sufficient for retribution but not culpability that compromises innocence—and a reason that the category is morally stable.

Murphy’s vicarious-punishment proposal can be read as accepting that adjustment. The burden is imposed on one person because of another’s wrong, but its justification is not identical to the ordinary desert of the recipient (Murphy, 2009). That may be coherent. It also concedes the paper’s narrower point: the account requires a special justification beyond ordinary retributive punishment.

6.2 Union, imputation, and an enlarged identity

A second route says that Christ’s union with humanity, covenantal headship, or incorporation of believers makes their guilt his in a deeper-than-fictional way. Woznicki develops an Edwardsean defense using a nonstandard account of personal identity and explicitly limits the project to conceptual coherence rather than claiming to have proved the account true or scripturally best (Woznicki, 2019). That modesty identifies exactly what such a defense can accomplish: if identity is constituted in the proposed way, the contradiction may be blocked.

Two questions then become decisive. First, does the relation preserve numerical and moral distinctions strongly enough for Christ to remain innocent? If “Christ and sinners are one” means that he is literally the culpable agent of their acts, innocence becomes difficult to state. If it means only representation or solidarity,

the responsibility gap reappears. Second, why does union transfer guilt rather than also immediately transfer innocence, virtue, or exemption in both directions? A morally selective identity relation needs explanation.

This is not a demand that every theological mystery be reduced to ordinary legal categories. It is a demand internal to the word *justice*. An exotic metaphysics can secure consistency, but its moral consequences cannot be inferred from the label “union” alone.

6.3 Communal and representative responsibility

A third route begins with the fact that people act through communities and inherit damaged practices they did not individually originate. Cawdron argues that communal sin can involve group non-agential moral responsibility and develops a communal-substitution model (Cawdron, 2024). This route has an advantage: it refuses the fiction that persons are isolated units whose only moral relations arise from their own discrete acts.

But communal responsibility normally distributes responsibility among participants in a shared practice. It does not obviously assign the community’s culpability to its least culpable member. Representation can authorize one person to act for a group, yet authority to represent is not identical to desert for every represented wrong. The account must specify whether Christ is a member, head, representative, victim, or constitutive bearer of the community, and which of these relations generates retributive liability.

The route may yield a powerful doctrine of solidarity: Christ enters the consequences of communal evil, exposes them, and acts reparatively on behalf of the community. That is already a substantial atonement claim. Its strength does not depend on describing the suffering as deserved punishment of Christ.

6.4 Penal consequences without punishment

A fourth route distinguishes punishment from “penal consequences.” Christ bears the divinely ordered consequences associated with wrongdoing, but the theory does not say that he is punished as a culpable wrongdoer. Jansen argues that this move fails to deliver the retributive justice such accounts promise: if Christ is not punished, then retribution has not been visited on the deserving agent; if the wrongdoer is not punished either, retributive desert remains unsatisfied (Jansen, 2025).

This criticism is well targeted. The penal-consequence route may explain costly satisfaction, deterrent display, solidarity, rescue, or restoration. It avoids unjustly censuring Christ precisely by removing a defining feature of punishment. What it cannot do without more is claim that retributive justice has been fulfilled through punishment. The reframing may be theologically preferable, but it should be named accurately.

Reply route	What it preserves	What it owes	Result under the test
Legal imputation	Liability and explicit penal language	Why assigned liability tracks desert rather than useful fiction	Vicarious punishment without recipient culpability
Union or enlarged identity	A deeper-than-contractual transfer of responsibility	A stable account of shared responsibility that preserves innocence	Identity strong enough for guilt may erase the innocence claim
Communal representation	Socially constituted wrongdoing and corporate solidarity	Why the representative deserves retribution for members’ wrongs	Costly solidarity or repair rather than punishment
Penal consequences	Serious, divinely ordered consequences borne for others	How retributive justice is fulfilled if no deserving wrongdoer is punished	Satisfaction, expression, or rescue rather than retribution

The point of the table is not that revision is dishonest. Theories often improve by abandoning an overstrong

description. The error is to use a nonretributive repair to defend the original claim that perfect retributive justice is satisfied by punishing an innocent substitute.

7. Context Raises the Burden; It Does Not Supply the Bridge

7.1 Plural images create underdetermination, not contradiction

Christian scripture describes the cross through sacrifice, ransom, reconciliation, victory, covenant, representation, example, cleansing, and participation. Substitutionary readings have real textual support; Gathercole, for example, defends substitution in Pauline material without thereby establishing every element of a later penal theory (Gathercole, 2015). The multiplicity of images does not prove contradiction. Different metaphors may describe complementary dimensions of one event.

The plurality does affect evidential burden. A defender of literal retributive transfer must show why penal categories govern the other images and why “for us” or “for our sins” entails punishment of an innocent person rather than representative, sacrificial, restorative, or liberative action. Conversely, a critic should not argue from plurality to the impossibility of a penal interpretation. The modest conclusion is underdetermination: the textual field does not make the disputed responsibility bridge self-evident.

This point also constrains appeals to Isaiah’s servant, sacrificial animals, or Pauline participation. These may ground substitution, representation, or sin-bearing. They do not by themselves answer whether the recipient of retributive punishment must be culpably responsible. Exegesis and moral analysis must meet rather than substitute for one another.

7.2 Historical development is context, not refutation

Substitutionary and sacrificial motifs long predate the Reformation. More precise penal formulations became especially prominent in Reformation and post-Reformation theology, where legal categories of guilt, satisfaction, and punishment were systematized (Craig, 2018a). This genealogy matters because it shows that one doctrinal grammar developed within a larger Christian field; it does not show that later formulations are false. Doctrines can become clearer over time.

The relevant evidential lesson is comparative. If the cross has been intelligibly interpreted through several models, literal retributive transfer cannot claim necessity merely because it is familiar. It must earn its explanatory priority by handling responsibility better than its rivals. Historical contingency weakens claims of obviousness, not claims of truth.

Cross-cultural comparison requires the same restraint. Many communities distinguish the guilty from the innocent, yet practices of collective liability, inherited status, sacrifice, and representative action vary widely. No uniform global norm can be assumed. The paper’s benchmark instead rests on the internal logic of retribution: punishment answers a person as culpably responsible. Cultural exceptions invite analysis of responsibility-making relations; they do not license arbitrary substitution.

7.3 A qualitative Bayesian comparison

Bayesian language can clarify without manufacturing numbers. Evidence favors a theory relative to alternatives when the evidence is less surprising on that theory and when the theory does not require costly auxiliary assumptions. Penal substitution fits some scriptural and devotional data well: it makes divine opposition to wrongdoing vivid, takes guilt seriously, and gives Christ’s suffering a direct “for us” structure. Those are genuine explanatory virtues.

The responsibility gap is an explanatory cost. The theory must add legal imputation, unusual identity conditions, communal metaphysics, or a revision of punishment. Alternatives that treat the cross as restorative

sacrifice, solidarity, participation, or the exposure of violence may fit the innocence of Christ more directly, though they face their own burdens concerning guilt, judgment, and why the cross is necessary or fitting. Without agreed priors or measurable likelihoods, numerical probabilities would create false precision. The justified conclusion is only comparative: an unexplained exception to the personal-responsibility condition lowers the literal retributive account's plausibility relative to otherwise adequate accounts that do not need that exception. Context can raise the cost of leaving the bridge unexplained; it cannot construct the bridge.

8. Atonement Without Transferred Retributive Desert

Rejecting transferred retributive desert does not make the cross morally or theologically empty. It changes the work the suffering is asked to do. Several neighboring accounts illustrate the range.

Stump develops a relational account in which atonement addresses the human will, shame, alienation, and union rather than reducing the cross to a penal exchange (Stump, 2018). Rutledge advances a restorative-sacrifice model that coordinates forgiveness, accountability, repair, and reconciliation (Rutledge, 2022). Heim construes the cross as entering and revealing sacrificial violence so that the scapegoat mechanism can be overcome (Heim, 2006). Communal and participatory accounts treat Christ as representative head or solidarity-bearing member without necessarily saying that he culpably deserves human punishment.

Restorative justice offers a useful moral comparison, provided it is not romanticized. Restorative practices seek acknowledgment, victim participation, repair, accountability, reintegration, and transformation. They do not imply that serious wrongs have no consequences or that reconciliation can be demanded from victims. Braithwaite's work on reintegrative shaming shows how condemnation can be directed at wrongdoing while the wrongdoer is invited back into community rather than fixed in a deviant identity (Braithwaite, 1989). The analogy suggests that justice has resources other than transferring penal pain.

These alternatives are not established merely because penal substitution faces a problem. Each must answer its own questions: why Christ's death rather than another act, how victims receive justice, whether wrongdoing is condemned, how agency is transformed, and what makes reconciliation more than unilateral amnesty. The point is logical space. Perfect justice may condemn, expose, repair, protect, restore, and transform without pretending that an innocent party committed the guilty party's act.

The cross can then be substitutionary in nonpenal senses. Christ may do for others what they cannot do, bear consequences caused by them, act as their representative, make costly repair, or enter the site of their violence to overcome it. Those claims require defense, but they do not commit the specific miscarriage identified here: calling suffering retributively deserved by a person stipulated to be innocent.

9. Conclusion

The strongest objection to literal retributive penal substitution is not a general aversion to sacrifice or an assumption that forgiveness must be costless. It is the personal-responsibility condition of just punishment. Retribution addresses a culpably responsible agent with the treatment that agent deserves. Christ is stipulated to be innocent of the wrongdoing in question. Consent makes sacrifice voluntary; it does not make the wrong his. Civil substitution transfers costs; it does not transfer penal censure. Scapegoat, scriptural, historical, and evidential considerations sharpen the concern but do not replace it.

Sophisticated defenses reveal the available paths. Legal imputation can assign liability, but must show why the assignment tracks desert. Union and identity accounts can make responsibility deeper than fiction, but must preserve innocence. Communal accounts can make agency socially constituted, but must explain why the representative deserves members' retribution. Penal-consequence accounts can avoid condemning Christ, but then no longer deliver punishment of the deserving wrongdoer.

The conclusion is correspondingly bounded. The ordinary literal account has not explained how punishment of

an innocent substitute satisfies retributive justice. That failure does not disprove Christian theism, the importance of the cross, substitution in every sense, or the logical possibility of an unusually strong metaphysics of union. It leaves a responsibility-bridge dilemma with three legitimate exits: defend a relation that creates desert without erasing innocence, reject a premise of the argument, or revise the event as nonretributive. Before transferred suffering is called just retributive punishment, one of those routes must be made explicit. Otherwise the language of punishment turns a voluntary act of solidarity, repair, or rescue into the very miscarriage it was meant to cure.

Formal Reconstruction of the Central Thesis

The central argument can first be stated as a plain-language syllogism:

1. If literal retributive penal substitution is just, Christ receives retributive punishment for human wrongdoing and has the desert that just retribution requires.
2. A person who did not culpably commit a wrong does not acquire retributive desert merely by accepting suffering or being declared liable; any representational exception needs an independently defended responsibility-making relation.
3. Christ did not culpably commit the human wrongdoing for which the theory says he is punished.
4. Therefore the ordinary responsibility condition, the ordinary nontransfer condition, Christ's innocence, and literal just retributive punishment cannot all be retained. A defender must reject one condition, reject the innocence claim, revise the event as nonretributive, or defend a principled responsibility bridge.

Symbols

Let the domain include persons and wrongs. Define:

- x : any person;
- w : any wrong;
- c : Christ;
- h : the human wrongdoing at issue;
- $P(x, w)$: person x is retributively punished for wrong w ;
- $J(x, w)$: that punishment of x for w is just;
- $C(c, h)$: Christ culpably committed the human wrongdoing at issue;
- $D(x, w)$: x has retributive desert for w ;
- $B(c, h)$: a defended bridge makes Christ responsible and deserving for h despite not culpably committing h .
- R : the ordinary responsibility condition;
- N : the ordinary nontransfer condition;
- I : Christ's stipulated innocence regarding h ;
- L : literal retributive penal substitution applied to c and h .

Symbolic form

The ordinary responsibility condition is:

$$R : \quad \forall x \forall w [(P(x, w) \wedge J(x, w)) \rightarrow D(x, w)].$$

The ordinary nontransfer condition is stated only for the target case. Unless a responsibility-making relation defeats it, Christ's noncommission of the human wrongdoing excludes desert for that wrongdoing:

$$N : \neg C(c, h) \rightarrow \neg D(c, h).$$

Christ's stipulated innocence regarding human wrongdoing is:

$$I : \neg C(c, h).$$

Literal retributive penal substitution asserts:

$$L : P(c, h) \wedge J(c, h).$$

From R and L , universal instantiation and modus ponens yield $D(c, h)$. From N and I , they yield $\neg D(c, h)$. Thus:

$$R, N, I, L \vdash \perp.$$

Response space

The defender has four structurally distinct responses:

1. deny or qualify R , allowing just retributive punishment without recipient desert;
2. deny or qualify I , making Christ a culpable committer of the relevant wrongdoing;
3. retain R , N , and I but replace L with a nonretributive claim—for example, voluntary suffering, satisfaction, representation, repair, or penal consequence;
4. defend $B(c, h)$ as a principled exception to N , showing how Christ can acquire genuine responsibility and desert without having culpably committed h .

The first response revises ordinary retributive justice; the second threatens a central innocence commitment; the third changes the theory under assessment; and the fourth carries the substantive burden identified throughout the paper.

Validity

Given R , N , I , and L , the contradiction follows by standard first-order steps: R and L yield $D(c, h)$, while N and I yield $\neg D(c, h)$. The derivation is therefore valid. A reply must reject, qualify, or reinterpret at least one premise; merely restating all four does not remove the contradiction.

Premise support and scope

The argument for R rests on retribution's connection to recipient desert; I and L state commitments of the target theory; and N expresses the target-specific nontransfer presumption defended through the consent, civil-analogy, and reply sections. The formal result is conditional. It does not prove $\neg B(c, h)$, universal nontransfer across every possible relation, or the impossibility of every coherent substitutionary account. It shows why repeating "imputation," "union," or "representation" is not enough: the bridge must be specified strongly enough to defeat N and carefully enough not to erase innocence.

References

- Braithwaite, J. (1989). *Crime, shame and reintegration*. Cambridge University Press. <https://doi.org/10.1017/CBO9780511804618>
- Cawdron, H. (2024). Communal sin, atonement, and group non-agential moral responsibility. *Religious Studies*, 60(4), 625–639. <https://doi.org/10.1017/S0034412523000665>
- Craig, W. L. (2018a). *The atonement*. Cambridge University Press. <https://doi.org/10.1017/9781108558020>
- Craig, W. L. (2018b). Is penal substitution incoherent? an examination of Mark Murphy’s criticisms. *Religious Studies*, 54(4), 509–526. <https://doi.org/10.1017/S003441251700018X>
- Duff, R. A. (2001). *Punishment, communication, and community*. Oxford University Press. <https://doi.org/10.1093/oso/9780195104295.001.0001>
- Feinberg, J. (1970). *Doing and deserving: Essays in the theory of responsibility*. Princeton University Press.
- Gathercole, S. (2015). *Defending substitution: An essay on atonement in Paul*. Baker Academic.
- Hart, H. L. A. (1968). *Punishment and responsibility: Essays in the philosophy of law*. Oxford University Press.
- Heim, S. M. (2006). *Saved from sacrifice: A theology of the cross*. William B. Eerdmans Publishing Company.
- Jansen, D. W. (2025). The failure of retributive justice in the penal consequence account of penal substitutionary atonement. *Philosophia Christi*, 27(1), 45–59. <https://doi.org/10.5840/pc20252715>
- Kyle, B. G. (2013). Punishing and atoning: A new critique of penal substitution. *International Journal for Philosophy of Religion*, 74(2), 201–218. <https://doi.org/10.1007/s11153-012-9382-1>
- Lewis, D. (1997). Do we believe in penal substitution? *Philosophical Papers*, 26(3), 203–209. <https://doi.org/10.1080/05568649709506566>
- Moore, M. S. (1997). *Placing blame: A theory of the criminal law*. Oxford University Press.
- Murphy, M. C. (2009). Not penal substitution but vicarious punishment. *Faith and Philosophy*, 26(3), 253–273. <https://doi.org/10.5840/faithphil200926314>
- Rutledge, J. C. (2022). *Forgiveness and atonement: Christ’s restorative sacrifice*. Routledge.
- Stott, J. R. W. (1986). *The cross of Christ*. InterVarsity Press.
- Stump, E. (2018). *Atonement*. Oxford University Press. <https://doi.org/10.1093/oso/9780198813866.001.0001>
- Woznicki, C. (2019). The coherence of penal substitution: An Edwardsean defence. *Tyndale Bulletin*, 70(1), 95–115. <https://doi.org/10.53751/001c.27714>