

Kontextuelle und globale Falsifikation wissenschaftlicher Modelle

Eine integrierte Theorie epistemischer Geltung

Abstract

Die klassische und häufig rezipierte Popper'sche Falsifikationstheorie operiert mit einem idealtypisch binären Bild wissenschaftlicher Rationalität, demzufolge eine Theorie durch eine widersprechende Beobachtung grundsätzlich in ihrer Geltung erschüttert wird (Popper [1934] 1959). Auch wenn Popper selbst methodologische Zusatzbedingungen berücksichtigt, bleibt diese binär-eliminative Struktur für die Analyse moderner Modellpraxis zu grob. Modelle wie die Newtonsche Mechanik, klassische Thermodynamik oder heutige Klimamodelle sind in Teilbereichen empirisch begrenzt, bleiben jedoch in bestimmten Domänen epistemisch unverzichtbar.

Dieses Paper entwickelt einen integrierenden Rahmen der Modellgeltung als wissenschaftstheoretische Spezialausarbeitung von Modellmanagement unter endlichen Bedingungen. Es verbindet Einsichten von Popper, Kuhn, Lakatos, da Costa/French und moderner Modelltheorie, unterscheidet globale und kontextuelle Falsifikation und führt den epistemischen Ermöglichungsraum $E(t)$ ein. $E(t)$ macht sichtbar, unter welchen methodischen, technischen und institutionellen Bedingungen Modelle formulierbar, prüfbar, stabilisierbar und durch Alternativen ersetzbar sind.

„Falsifikation“ wird dabei nicht als rein logisches Wahrheitskriterium verstanden, sondern als rational-pragmatischer Mechanismus der Modellbewertung. Entscheidend ist nicht nur, ob die Leistung eines Modells in einer Domäne sinkt, sondern ob verfügbare Alternativen innerhalb derselben relevanten Domäne über Zeit stabil höhere epistemische Konkurrenzfähigkeit zeigen. Die vorgeschlagene rekonstruktive Nutzenstruktur verbindet Approximate Truth, explanatorische Kraft und Modellkosten, ohne wissenschaftliche Rationalität als allgemein algorithmisch berechenbar zu behandeln. Eine Fallstudie zu Klimamodellen zeigt exemplarisch, wie Ensemble-Methoden, Parametrisierungen und probabilistische Gewichtungungsverfahren zur Domänenpräzisierung beitragen können. Die Theorie rekonstruiert damit zentrale Strukturen moderner Modellpraxis und erklärt, warum Modelle trotz Teilfalsifikationen stabil bleiben, wann sie revidiert werden und unter welchen Bedingungen globale Modelleeliminierung erfolgt.

Keywords

kontextuelle Falsifikation; globale Falsifikation; Modellgeltung; Modellwahl; Approximate Truth; Wahrheitsnähe; wissenschaftliche Modelle; epistemische Kosten; Modellrevision; epistemischer Ermöglichungsraum; Wissenschaftstheorie; wissenschaftlicher Realismus; Simulationsmodelle; Klimamodelle; Epistemik

Stefan Rapp

Independent Researcher

Last revised: 27 April 2026

<https://orcid.org/0009-0004-0847-9164>

<https://doi.org/10.5281/zenodo.17692470>

Epistemics Project: <https://www.epistemics.de>

© 2026 Stefan Rapp — Licensed under CC BY-NC-ND 4.0

Inhaltsverzeichnis

1. Von binärer Falsifikation zu domänenrelativer Modellgeltung.....	3
2. Forschungsstand.....	5
2.1 Popper: Falsifikation als binäres Ideal.....	5
2.2 Duhem-Quine: Unterbestimmtheit empirischer Tests.....	5
2.3 Kuhn: Anomalien ohne Eliminierung.....	5
2.4 Lakatos: Forschungsprogramme.....	5
2.5 Modelltheorie.....	6
2.6 Approximate Truth.....	6
2.7 Forschungslücke.....	7
2.8 Eigenbeitrag des Papers.....	8
3. Grundbegriffe.....	9
4. Kriterien moderner Modellbewertung im vorgeschlagenen Rahmen.....	15
5. Proposition: Zweistufige Falsifikationsstruktur.....	17
5.1 Begrenzung von Immunisierungswegen.....	18
6. Wann Modelle verschwinden.....	19
7. Warum Modelle bleiben.....	20
7.1 Robustheit durch Modellfamilien.....	20
7.2 Nutzungskontexte.....	20
8. Fallstudie: Klimamodelle.....	21
8.1 Klimamodelle als Ensemble-Struktur.....	21
8.2 Parametrisierung als epistemische Operation.....	21
8.3 Bayesian Updating und adaptive Modellierung.....	22
8.4 Domänenstruktur.....	23
8.5 Miniaturisierte formale Illustration.....	24
8.6 Konsequenzen für den epistemischen Status.....	24
9. Approximate Truth.....	25
10. Warum eine integrierte Theorie bisher fehlte.....	27
10.1 Historische Gründe.....	27
10.2 Logische Gründe.....	27
10.3 Institutionelle Gründe.....	27
11. Epistemischer Ermöglichungsraum.....	28
11.1 Dynamik des epistemischen Ermöglichungsraums.....	28
11.2 Struktur des epistemischen Ermöglichungsraums $E(t)$	29
12. Entscheidungslogik der Modellwahl.....	30
12.1 Struktur der Nutzenfunktion.....	30
12.2 Skalierung und empirische Verankerung der Nutzenfunktion.....	33
12.3 Diagnoseprotokoll kontextueller und globaler Falsifikation.....	34
13. Verhältnis zum wissenschaftlichen Realismus.....	36
14. Falsifikation als domänenrelative Geltungsrevision.....	37
Lokale wissenschaftstheoretische Begriffsfassung dieses Papers.....	38
Literaturverzeichnis.....	41
Appendix: Didaktische Beispiele zur zweistufigen Falsifikationsstruktur.....	43

1. Von binärer Falsifikation zu domänenrelativer Modellgeltung

Die Falsifikationstheorie nach Popper ([1934] 1959) prägte über Jahrzehnte das öffentliche und akademische Bild wissenschaftlicher Rationalität. In ihrer klassischen und häufig rezipierten Form erscheint Falsifikation als idealtypisch binärer Test theoretischer Geltung: Eine Theorie gerät in ihrer wissenschaftlichen Akzeptabilität unter Druck, wenn eine ihrer Vorhersagen empirisch scheitert. Die Praxis der Natur- und Sozialwissenschaften zeigt jedoch ein komplexeres Bild. Viele Modelle sind idealisiert, nur partiell repräsentativ, empirisch begrenzt in Subdomänen und dennoch epistemisch stabil. Dieses Paper entwickelt daher eine integrierte Theorie wissenschaftlicher Modellgeltung, die zentrale Einsichten aus Popper, Kuhn, Lakatos, da Costa/French sowie der modernen Modelltheorie in einer gemeinsamen Struktur rationaler Modellbewertung zusammenführt.

Beispiele hierfür sind die Newtonsche Mechanik, die bei relativistischen Geschwindigkeiten versagt (Einstein 1905), klassische Thermodynamik, die mikroskopisch unpräzise ist, oder Klimamodelle, die kontinuierlich rekaliert werden (Oreskes et al. 1994). Gleichzeitig gibt es Modelle, die vollständig verschwunden sind, wie die Phlogiston-Theorie oder der klassische Äther.

Das Kernproblem lautet: Unter welchen Bedingungen führt Falsifikation zur Eliminierung eines Modells – und wann bleibt ein Modell trotz Teilfalsifikation epistemisch stabil?

Dieses Problem wird im Folgenden als Spezialfall der Epistemik verstanden, also als Spezialfall von Modellmanagement unter endlichen Bedingungen (Rapp 2026a). Das Paper ist damit als Modul innerhalb eines umfassenderen Forschungsprogramms zu verstehen, in dem Friktion, Domänenstruktur, Suche und Revision in eigenständigen Arbeiten ausgearbeitet werden. Die vorliegende Studie überträgt diese Architektur gezielt auf den Spezialfall wissenschaftlicher Falsifikation. Wissenschaftliche Modelle besitzen keine unbegrenzte Geltung, sondern werden innerhalb bestimmter Domänen stabilisiert, geprüft, eingeschränkt und revidiert. Die allgemeinen Begriffe Modell, Geltung, Domäne, Kosten und Revision werden in diesem Paper daher auf den Spezialfall wissenschaftlicher Falsifikation angewandt.

Falsifikation erscheint in diesem Rahmen nicht primär als einfache Eliminationsoperation, sondern als spezifische Form epistemischer Friktion im Sinne eines Grenzsymbols endlicher Tragfähigkeit (Rapp 2026b): Sie zeigt an, dass die bisherige Stabilisierung eines Modells in einer bestimmten Domäne nicht mehr ohne Weiteres tragfähig ist. Friktion fungiert dabei als prozeduraler Indikator einer belasteten Nutzenstruktur. Sie tritt auf, wenn empirische Abweichungen $AT(M, D, t)$ schwächen, erklärende Defizite $EK(M, D, t)$ reduzieren, Modellkosten $C(M, D, t)$ steigen oder verfügbare Alternativen M^* relativ konkurrenzfähiger werden. Ob daraus eine lokale Revision, eine Einschränkung des Geltungsbereichs oder eine globale Modellablösung folgt, hängt von der stabilen relativen Konkurrenzfähigkeit des Modells innerhalb von $E(t)$ ab.

Dieses Paper entwickelt einen integrierten theoretischen Rahmen, der diese Frage explizit und in einem einheitlichen Modell rekonstruiert. Der Ansatz unterscheidet globale und kontextuelle Falsifikation, beschreibt den epistemischen Ermöglichungsraum $E(t)$ und integriert Approximate Truth als graduelle, domänenspezifische Bewertungsgröße. Falsifikation wird dabei nicht als rein logische Wahrheitsoperation verstanden, sondern als rational rekonstruierbare Veränderung von Modellgeltung innerhalb eines realen wissenschaftlichen Ermöglichungsraums. Entscheidend ist nicht der isolierte Leistungsabfall eines Modells, sondern seine relative Konkurrenzfähigkeit gegenüber verfügbaren Alternativen innerhalb derselben relevanten Domäne. Die Theorie ist primär deskriptiv und rekonstruktiv: Sie rekonstruiert die tatsächliche Modellpraxis moderner Wissenschaften und wird exemplarisch anhand von Klimamodellen demonstriert.

Die im Folgenden eingeführte Nutzenfunktion $U(M, D, t)$ ist primär rekonstruktiv zu verstehen. Sie beschreibt keine allgemeine Berechnungsmaschine wissenschaftlicher Rationalität, sondern macht explizit, wie Modellwahl in realen Wissenschaften rekonstruiert werden kann, wenn Approximate Truth, explanatorische Kraft und Modellkosten gegeneinander abgewogen werden. Innerhalb dieses Papers formalisiert $U(M, D, t)$ die domänenspezifische Geltung wissenschaftlicher Modelle unter endlichen Bedingungen. Zugleich besitzt dieselbe Struktur eine schwach normative Lesbarkeit, da sie sichtbar macht, unter welchen Bedingungen Modellentscheidungen als rational nachvollziehbar gelten können. Damit ist die Theorie sowohl eine deskriptive Rekonstruktion der Modellpraxis als auch ein Rahmen epistemischer Optimalität, ohne daraus eine universelle Entscheidungsformel abzuleiten.

2. Forschungsstand

2.1 Popper: Falsifikation als binäres Ideal

Popper ([1934] 1959) formuliert Falsifikation als methodologisches Gegenmodell zu Verifikation: Wissenschaftliche Theorien müssen so beschaffen sein, dass sie an möglichen Beobachtungen scheitern können. In der klassischen und häufig rezipierten Popper'schen Struktur erscheint Falsifikation dabei als idealtypisch binärer Test theoretischer Geltung. Diese Struktur setzt starke Bedingungen voraus: relativ globale Geltungsansprüche, eindeutige Theorie-Beobachtung-Zuordnungen und hinreichend klar isolierbare Prüfbedingungen. Gerade idealisierte, approximative und simulierte Modelle passen nur begrenzt zu dieser binären Struktur. Die Spannung zwischen Poppers Ideal und realer Modellpraxis ist in der Literatur vielfach beschrieben worden (Hacking 1983; Cartwright 1983), wurde jedoch selten in eine systematische Theorie kontextueller Falsifikation überführt.

2.2 Duhem-Quine: Unterbestimmtheit empirischer Tests

Die Duhem-Quine-Problematik zeigt, dass empirische Tests selten isolierte Theorien betreffen, sondern Bündel aus Modellannahmen, Hilfhypothesen, Messpraktiken und Hintergrundbedingungen. Der hier entwickelte Rahmen übernimmt diese Einsicht, rekonstruiert sie jedoch domänenrelativ: Entscheidend ist nicht nur, welche Annahme durch eine Abweichung betroffen ist, sondern ob daraus lokale Revision, Domänenbegrenzung oder globale Modellablösung folgt. Falsifikation erscheint damit nicht als unmittelbare logische Negation eines Einzelmodells, sondern als strukturierte Veränderung von Modellgeltung innerhalb bestimmter Prüf- und Anwendungsbedingungen.

2.3 Kuhn: Anomalien ohne Eliminierung

Kuhn (1962) zeigt, dass wissenschaftliche Paradigmen Anomalien überstehen. Er beschreibt jedoch keine fein strukturierte Theorie modellbezogener Falsifikation. Der Mangel besteht insbesondere darin, dass Kuhn nicht spezifiziert, wie sich Anomalien über Subdomänen verteilen und warum sie manche Modelle kaum beeinträchtigen, andere aber untergraben. Die Literatur betont zwar Kuhns historische und soziologische Perspektive, doch fehlt eine formal präzise Modelltheorie des Umgangs mit Subdomänen (vgl. Lakatos 1970; Weisberg 2013).

2.4 Lakatos: Forschungsprogramme

Lakatos (1970) erklärt Stabilität über einen „harten Kern“, ohne jedoch zu bestimmen, wann Modelle tatsächlich verschwinden. Zudem bleibt unklar, wie konkurrierende Programme im Detail gegeneinander abgewogen werden sollen, insbesondere wenn beide partiell, aber nicht global, falsifiziert sind. Gerade hier zeigt sich, dass Lakatos' Programmstruktur kein formales Instrument zur Bewertung domänenspezifischer Approximate Truth bereitstellt.

2.5 Modelltheorie

Modelltheoretische Ansätze (Cartwright 1983; Weisberg 2013; Morgan und Morrison 1999) betonen die idealisierte und domänenspezifische Natur wissenschaftlicher Modelle, liefern aber keine Theorie, wie Falsifikation in solchen Strukturen wirkt oder wie Modelle systematisch gegeneinander abgewogen werden. Ergänzend zeigen Erklärungstheorien, dass explanatorische Kraft unterschiedlich rekonstruiert werden kann: etwa als Vereinheitlichung vieler Phänomene unter wenige Erklärungsmuster (Kitcher 1981, 1989) oder als kausal-interventionelle Aufschlusskraft über mögliche Veränderungen (Woodward 2003). Ähnliche Einsichten wurden auch in Debatten über Modellpluralismus, Idealisierung, Robustheit und Simulation entwickelt. Arbeiten zu pluralistischer Wissenschaftspraxis, zu isolierenden Modellierungsstrategien und zu simulationsbasiertem Wissen zeigen, dass wissenschaftliche Modelle häufig nicht als globale Abbilder, sondern als selektive, domänengebundene und zweckabhängige Strukturen fungieren. Der Eigenbeitrag des vorliegenden Papers besteht daher nicht darin, die Domänengebundenheit wissenschaftlicher Modelle erstmals zu behaupten, sondern darin, diese Einsicht mit einer zweistufigen Theorie kontextueller und globaler Falsifikation, einer rekonstruktiven Nutzenstruktur und dem epistemischen Ermöglichungsraum $E(t)$ zu verbinden.

2.6 Approximate Truth

Niiniluoto (1987, 1998) und Oddie (1986) bieten graduelle Wahrheitskonzepte, erklären aber nicht, warum manche Modelle trotz Teilfalsifikationen weiterbestehen, während andere vollständig verschwinden. Besonders fehlt die Kopplung von Approximate Truth an Nutzungskontexte, Modellkosten und institutionelle Stabilisierung. Damit bleibt unklar, wie graduelle Wahrheitsgrade in realen Modellportfolios operativ wirken sollen.

2.7 Forschungslücke

Die einschlägige Forschung hat bereits zahlreiche Elemente bereitgestellt, die für eine Theorie moderner Modellbewertung zentral sind: die Unterbestimmtheit empirischer Tests, die Persistenz von Anomalien, die Stabilität von Forschungsprogrammen, die Domänengebundenheit wissenschaftlicher Modelle, graduelle Wahrheitsnähe, Modellpluralismus, Robustheitsanalysen und Verfahren statistischer Modellselektion. Was jedoch weitgehend fehlt, ist ein integrierter Rahmen, der diese Einsichten ausdrücklich auf die Unterscheidung zwischen kontextueller und globaler Falsifikation bezieht. Die Forschungslücke liegt daher nicht in der Behauptung, dass Modelle domänenspezifisch, idealisiert oder plural organisiert sind. Sie liegt in der systematischen Kopplung von domänenrelativer Modellgeltung, relativer epistemischer Konkurrenzfähigkeit, Modellkosten, verfügbaren Alternativen und dem epistemischen Ermöglicungsraum $E(t)$. Ein solcher Rahmen muss zugleich erklären:

- wann Falsifikation global wirkt,
- wann sie kontextuell bleibt,
- wie Approximate Truth innerhalb spezifischer Domänen funktioniert,
- wie Modellkosten und explanatorische Kraft in Modellentscheidungen eingehen,
- wie verfügbare Alternativen die relative Konkurrenzfähigkeit eines Modells bestimmen,
- wie der epistemische Ermöglicungsraum $E(t)$ die Menge realer Modellalternativen begrenzt und ermöglicht.

Der hier vorgeschlagene Rahmen setzt genau an dieser Lücke an. Er behandelt Falsifikation nicht als isoliertes logisches Ereignis, sondern als Veränderung der relativen Modellgeltung innerhalb einer Domäne. Dadurch werden Anomalien, Modellpersistenz, Approximate Truth, Modellkosten und verfügbare Alternativen nicht getrennt betrachtet, sondern als miteinander verbundene Dimensionen wissenschaftlicher Modellbewertung rekonstruiert. Dies ist insbesondere für datenintensive, simulationsbasierte Wissenschaften relevant, in denen kontextuelle Falsifikation und Modellportfolios zur Standardpraxis geworden sind.

2.8 Eigenbeitrag des Papers

Das vorliegende Paper leistet vier miteinander verbundene Beiträge zur Wissenschaftstheorie und Modellphilosophie. Es versteht sich als lokale wissenschaftstheoretische Spezialarchitektur innerhalb eines allgemeineren Verständnisses von Modellmanagement unter endlichen Bedingungen. Sein Ziel besteht nicht darin, eine unabhängige Metatheorie neben diesem Rahmen zu entwickeln, sondern die allgemeinen Begriffe von Modellgeltung, Domänenbindung, Kosten, Friktion und Revision für den Spezialfall wissenschaftlicher Falsifikation formal anschlussfähig zu machen.

1. Der Begriff des epistemischen Ermöglichungsraums $E(t)$.

$E(t)$ wird als dynamische Struktur methodischer, technischer und institutioneller Bedingungen expliziert, die bestimmt, welche Modelle formulierbar, evaluierbar und stabilisierbar sind. Damit wird eine Dimension präzise modelliert, die in der klassischen Falsifikationstheorie Poppers nur implizit oder gar nicht berücksichtigt wird.

2. Die zweistufige Falsifikationsstruktur.

Das Paper unterscheidet strikt zwischen kontextueller und globaler Falsifikation. Globale Falsifikation wird nicht mehr als binärer Wahrheitsbruch verstanden, sondern als Zustand, in dem ein Modell unter den Bedingungen von $E(t)$ in keiner relevanten Domäne D epistemisch konkurrenzfähig bleibt. Diese Struktur rekonstruiert die Stabilität partiell falsifizierter Modelle systematisch.

3. Die integrierte Nutzenfunktion $U(M, D, t)$.

Mit der Nutzenfunktion

$$U(M, D, t) = \alpha \cdot AT(M, D, t) + \beta \cdot EK(M, D, t) - \gamma \cdot C(M, D, t)$$

wird ein Rahmen vorgeschlagen, der Approximate Truth, explanatorische Kraft und Modellkosten in einer einzigen Bewertungsstruktur zusammenführt und explizit an $E(t)$ bindet. Die Funktion ist so formuliert, dass sie mit etablierten Gütemaßen und Selektionskriterien (Fit-Metriken, Informationskriterien, Komplexitätsmaßen) der einzelnen Disziplinen kompatibel gehalten werden kann.

4. Die Anwendung auf simulationsbasierte Modellportfolios.

Am Beispiel von Klimamodellen wird gezeigt, wie Ensemble-Strukturen, Parametrisierungen und adaptive Gewichtungsverfahren innerhalb des vorgeschlagenen Rahmens als Fälle kontextueller Falsifikation und domänenspezifischer Modellrevision rekonstruiert werden können. Dadurch wird die Theorie unmittelbar an eine zentrale Klasse moderner, daten- und simulationsintensiver Modellpraktiken anschlussfähig, ohne zu behaupten, dass alle Klimamodelle selbst einem einheitlichen bayesianischen Aktualisierungsmechanismus folgen.

3. Grundbegriffe

Die folgenden Definitionen präzisieren die für dieses Paper zentralen Begriffe. Sie beziehen allgemeine Kategorien des Modellmanagements unter endlichen Bedingungen auf den Spezialfall wissenschaftlicher Falsifikation. Modellgeltung, Domänenbindung, Kosten, Friktion und Revision werden dadurch nicht als abstrakte Hintergrundbegriffe belassen, sondern in eine formal anschlussfähige Struktur überführt. Die Größen $U(M, D, t)$, $AT(M, D, t)$, $C(M, D, t)$ und $E(t)$ dienen dazu, diese Struktur im Kontext wissenschaftlicher Modellbewertung explizit rekonstruierbar zu machen.

Notation (Übersicht)

M – Modell

M^* – verfügbares Alternativmodell

$D(M)$ – Geltungsbereich eines Modells

$D_1, D_2, \dots, D_n$ – Subdomänen des Geltungsbereichs

$AT(M, D, t)$ – Approximate Truth des Modells in Domäne D zum Zeitpunkt t

$EK(M, D, t)$ – explanatorische Kraft des Modells in Domäne D zum Zeitpunkt t

$C(M, D, t)$ – Modellkosten des Modells in Domäne D zum Zeitpunkt t

$U(M, D, t)$ – domänenspezifische Nutzenfunktion

$\Delta U(M^*, M, D, t)$ – relative Nutzendifferenz zwischen Alternativmodell M^* und Modell M

$E(t)$ – epistemischer Ermöglichungsraum zum Zeitpunkt t

$R(M, D_i)$ – Betrachtung oder Anwendung eines Modells innerhalb einer Subdomäne D_i

ε – kontextabhängiger Toleranzbereich epistemischer Konkurrenzfähigkeit

$D(M)'$ – verbleibender Geltungsbereich nach Einschränkung oder Ausschluss einer problematischen Subdomäne

$Rev(M, D_i)$ – Revision eines Modells oder Teilmodells innerhalb einer problematischen Subdomäne

Definition 1: Modell (M)

Ein Modell ist eine idealisierte mathematische oder konzeptuelle Struktur zur Repräsentation definierter Aspekte eines Zielsystems unter bestimmten Bedingungen (Cartwright 1999).

Definition 2: Geltungsbereich $D(M)$

Der Geltungsbereich eines Modells bezeichnet jene Domäne, innerhalb derer ein Modell zuverlässige, erklärkräftige oder funktional tragfähige Ergebnisse liefert. Eine Domäne ist dabei nicht bloß ein äußerer Anwendungsbereich, sondern ein Bereich stabiler Geltungsbedingungen: Innerhalb dieser Bedingungen kann ein Modell geprüft, verglichen, eingeschränkt oder gegenüber Alternativen verteidigt werden. $D(M)$ bezeichnet somit nicht die globale Wahrheit eines Modells, sondern seine domänenrelative epistemische Geltung. Diese domänenrelative Lesart schließt an die Theorie von Domänen, Grenzen und Übergangsfunktionen an, in der Geltung als abhängig von den jeweiligen Stabilitätsbedingungen einer Domäne rekonstruiert wird (Rapp 2026e).

Definition 3: Subdomänen D_1, D_2

$D(M)$ kann in eine Menge von Subdomänen $\{D_1, D_2, \dots, D_n\}$ zerlegt werden. Diese Zerlegung ist häufig Ergebnis wissenschaftlicher Revision, Domänenpräzisierung oder Modellvergleichs. Für einfache Fälle kann zwischen erhaltenen Domänen und problematischen Domänen unterschieden werden: Erhaltene Domänen bezeichnen jene Bereiche, in denen M epistemisch optimal oder konkurrenzfähig bleibt; problematische Domänen bezeichnen jene Bereiche, in denen M geschwächt, revisionsbedürftig oder gegenüber Alternativen nicht mehr epistemisch konkurrenzfähig ist. Die binäre Unterscheidung zwischen D_1 und D_2 ist daher nur ein vereinfachter Spezialfall einer allgemeineren Subdomänenstruktur.

Definition 4: Epistemische Optimalität

Ein Modell ist in einer Domäne D epistemisch optimal, wenn es bezogen auf alle im epistemischen Ermöglichungsraum $E(t)$ verfügbaren Alternativen das günstigste Verhältnis von Approximate Truth, explanatorischer Kraft und Modellkosten aufweist. Epistemische Optimalität ist daher kein absoluter Zustand eines Modells, sondern eine relationale Bewertung: Ein Modell ist optimal nur relativ zu einer Domäne, zu einem Zeitpunkt t und zu den tatsächlich verfügbaren Alternativen innerhalb von $E(t)$. Operativ wird diese Bewertung durch die domänenspezifische Nutzenfunktion $U(M, D, t)$ rekonstruiert.

Approximate Truth bezeichnet dabei die domänenspezifische Passung eines Modells zu relevanten empirischen, strukturellen oder interventionell prüfbar Merkmalen. Explanatorische Kraft bezeichnet demgegenüber die erklärende Ordnungs- und Erschließungsleistung eines Modells: seine Fähigkeit, Phänomene verständlich zu machen, Zusammenhänge zu strukturieren, kausale oder gegenfaktische Abhängigkeiten sichtbar zu machen und eine Domäne theoretisch zu orientieren. Damit wird zwischen Passung und Erklärung analytisch unterschieden: Ein Modell kann gut zu bestimmten Daten oder Strukturen passen, ohne bereits hohe Erklärungskraft zu besitzen; umgekehrt bleibt Erklärungskraft epistemisch begrenzt, wenn die Passung zu den relevanten Merkmalen einer Domäne zu schwach ist.

Modellkosten $C(M, D, t)$ umfassen drei Dimensionen:

- (1) epistemische Kosten (z. B. Fehleranfälligkeit, Sensitivität, Unsicherheitsbreite von Vorhersagen, Interpretierbarkeit/Transparenz, Fit-Qualität),
- (2) technische Kosten (Rechenaufwand, Datenbedarf, Implementierungskomplexität),
- (3) institutionelle Kosten (Normierungsgrad, Verfügbarkeit von Infrastrukturen).

Einige dieser Kostenkomponenten sind global (z. B. grundlegende Implementierungs- oder Infrastrukturkosten), andere sind domänenspezifisch (z. B. Datenanforderungen oder numerische Stabilität in bestimmten Domänen). Für die Nutzenfunktion wird dies in $C(M, D, t)$ gebündelt.

Definition 5: Kontextuelle Falsifikation

Eine Beobachtung O führt zu kontextueller Falsifikation, wenn sie zeigt, dass eine Subdomäne D_2 existiert, in der M nicht länger epistemisch akzeptabel oder konkurrenzfähig ist, während mindestens eine Subdomäne D_1 erhalten bleibt, in der M epistemisch optimal oder konkurrenzfähig bleibt. Kontextuelle Falsifikation ist damit eine spezifische Form epistemischer Friktion: Sie zeigt die Grenze der bisherigen Modellgeltung an, ohne das Modell als Ganzes zu eliminieren. Ihre typische Folge ist nicht globale Verwerfung, sondern Revision, Einschränkung oder Neuordnung des Geltungsbereichs. Eine solche Neuordnung ist jedoch nur dann epistemisch legitim, wenn sie mehr leistet als die bloße Rettung eines geschwächten Modells: Sie muss neue prüfbare Unterschiede sichtbar machen, Fehlerquellen diagnostisch trennen, Modellanwendungen präzisieren oder neue Erklärungs-, Prognose- oder Interventionsleistungen ermöglichen. In diesem Sinn ist kontextuelle Falsifikation eng mit Revision unter endlichen Bedingungen verbunden, insofern sie Modellgeltung nicht einfach negiert, sondern begründete Anpassungen, Einschränkungen oder Transformationen des Geltungsbereichs auslöst (Rapp 2026c).

Definition 6: Globale Falsifikation

Ein Modell M ist zu einem Zeitpunkt t global falsifiziert, wenn es unter den Bedingungen des epistemischen Ermöglichungsraums $E(t)$ in keiner relevanten Domäne D epistemisch optimal oder epistemisch konkurrenzfähig bleibt. Epistemische Optimalität liegt in einer Domäne D vor, wenn $U(M, D, t)$ größer oder gleich dem Nutzenwert aller im Ermöglichungsraum $E(t)$ verfügbaren Alternativmodelle ist. Epistemische Konkurrenzfähigkeit liegt vor, wenn $U(M, D, t)$ höchstens um einen kontextabhängigen Toleranzbereich ε unter dem Nutzenwert der jeweils besten Alternative $U(M^*, D, t)$ liegt.

Formal lässt sich globale Falsifikation so fassen: Für alle relevanten Domänen D existiert mindestens ein Alternativmodell M^* , sodass gilt:

$$U(M, D, t) + \varepsilon < U(M^*, D, t).$$

Der Toleranzparameter ε reflektiert, dass Modelle mit geringfügig schlechteren Nutzenwerten in der Praxis weiterhin konkurrenzfähig bleiben können. Damit ε nicht zur Immunisierung eines geschwächten Modells wird, ist der Toleranzbereich nicht beliebig zu setzen, sondern aus den jeweiligen disziplinären Bewertungspraktiken, Fehlertoleranzen, Unsicherheitsmaßen und Anwendungskontexten zu rekonstruieren. Operativ kippt ε dann von legitimer Toleranz in Immunisierung, wenn er größer wird als die in der jeweiligen Domäne anerkannten Fehler- und Unsicherheitsmargen. Liegt die Nutzendifferenz $\Delta U(M^*, M, D, t)$ systematisch oberhalb dieser disziplinären Standards, kann sie nicht mehr durch ε absorbiert werden, ohne dass die Bewertungspraktik der Disziplin selbst außer Kraft gesetzt wird. ε bezeichnet damit nicht eine schützende Zone für unterlegene Modelle, sondern den Bereich, innerhalb dessen Konkurrenzfähigkeit noch von domänenüblichen Bewertungsverfahren als gegeben anerkannt wird. Globale Falsifikation liegt daher nicht bereits bei bloßer oder kurzfristiger Unterlegenheit gegenüber einer Alternative vor, sondern erst bei stabiler relativer Nicht-Konkurrenzfähigkeit in allen relevanten Domänen, also dann, wenn die Unterlegenheit gegenüber verfügbaren Alternativen nicht nur punktuell, sondern über die jeweils maßgeblichen Prüf- und Anwendungskontexte hinweg bestehen bleibt.

Ergänzung: Relevante Domäne

Als relevante Domäne gilt ein wissenschaftlicher Nutzungsbereich, in dem ein Modell weiterhin erklärende, prognostische, interventionelle oder strukturierende Leistung innerhalb aktueller Forschungs- oder Anwendungspraxis erbringt. Nicht jede Restnutzung eines Modells begründet daher epistemische Konkurrenzfähigkeit. Historische, didaktische, metaphorische oder rein illustrative Verwendungen können weiterhin legitim sein, zählen aber nicht als relevante Domänen im Sinne globaler Falsifikation. Ein Modell kann also historisch oder didaktisch erhalten bleiben und dennoch epistemisch global falsifiziert sein, wenn es in keiner aktuellen wissenschaftlichen Domäne mehr konkurrenzfähig ist. Die Relevanz einer Domäne wird dabei nicht allein aus dem Nutzenwert $U(M, D, t)$ abgeleitet, sondern aus der bestehenden wissenschaftlichen Praxis, ihren Zielgrößen, Prüfverfahren, Anwendungskontexten und anerkannten Bewertungsordnungen rekonstruiert. Dadurch wird vermieden, dass Domänenrelevanz bloß aus der Leistungsbewertung des jeweils untersuchten Modells folgt.

Definition 7: Epistemischer Ermöglichungsraum E(t)

Der Raum methodischer, mathematischer, institutioneller und technischer Voraussetzungen, der bestimmt:

- welche Modelle formulierbar sind,
- welche Daten verfügbar sind,
- welche Idealisierungen zulässig sind,
- welche Modelle stabilisiert werden können,
- wie Modellkosten $C(M, D, t)$ überhaupt bestimmbar werden.

$E(t)$ ist dynamisch: Technische Innovationen, Datenverfügbarkeit und institutionelle Normen verändern den Raum der möglichen Modelle.

Definition 8: Approximate Truth

Approximate Truth bezeichnet den Grad stabilisierter Passung zwischen Modellstruktur und den innerhalb einer Domäne relevanten Vergleichsmerkmalen. Zu diesen Vergleichsmerkmalen können Beobachtungsdaten, Messstrukturen, robuste Systemmerkmale, Interventionsresultate oder theoretisch ausgezeichnete Strukturen gehören. Approximate Truth ist damit keine globale Eigenschaft eines Modells und kein unmittelbarer Zugriff auf ein Zielsystem an sich, sondern eine domänenspezifische und rekonstruktive Bewertungsgröße.

Formal lässt sich Approximate Truth als Ähnlichkeits- und Passungsmetrik fassen:

$$AT(M, D, t) = \sum_i w_i \cdot \text{sim}(M, S_i, D, t).$$

Dabei bezeichnet S_i ein innerhalb der Domäne D als relevant bestimmtes Vergleichsmerkmal; w_i beschreibt dessen relative Gewichtung innerhalb der jeweiligen wissenschaftlichen Bewertungsordnung. Die Interventionspraxis kann als prüfbarer Passungsbezug eingehen, etwa wenn ein Modell stabile Veränderungen unter Eingriffen korrekt abbildet. Die theoretische Deutung, kausale Einordnung oder gegenfaktische Erschließung solcher Zusammenhänge gehört jedoch zur explanatorischen Kraft $EK(M, D, t)$ und bleibt von AT analytisch zu unterscheiden.

In der praktischen Modellbewertung wird Approximate Truth stets zusammen mit Unsicherheitsabschätzungen der Modellvorhersagen betrachtet. Hohe Vorhersageunsicherheit reduziert den effektiven epistemischen Wert eines Modells, selbst wenn der mittlere Fit hoch ist, und kann daher sowohl $AT(M, D, t)$ als auch die epistemischen Kosten $C(M, D, t)$ beeinflussen.

Definition 9: Domänenoperatoren

Die folgenden Operatoren beschreiben unterschiedliche Formen der Veränderung oder Präzisierung von Modellgeltung.

- Restriktion: $R(M, D_i) = M|D_i$ bezeichnet die Anwendung oder Betrachtung eines Modells innerhalb einer bestimmten Subdomäne D_i .
- Domänenendifferenz: $D(M)' = D(M) \setminus D_i$ bezeichnet den verbleibenden Geltungsbereich eines Modells nach Ausschluss oder Aufgabe einer problematischen Subdomäne D_i .
- Revision: $Rev(M, D_i)$ bezeichnet die gezielte Anpassung eines Modells oder Teilmodells innerhalb einer problematischen Subdomäne D_i .
- Zerlegung: $Z(M) = \{D_1, D_2, \dots, D_n\}$ bezeichnet die Aufteilung des Geltungsbereichs eines Modells in mehrere prüf- und bewertbare Subdomänen.

Diese Operatoren unterscheiden drei Fälle, die in der Modellpraxis oft vermischt werden: die bloße Betrachtung eines Modells innerhalb einer Subdomäne, die Einschränkung seines Geltungsbereichs und die lokale Revision innerhalb einer problematischen Domäne.

Zur zeitlichen Dynamik

Da sich der epistemische Ermöglichungsraum $E(t)$ im Verlauf der wissenschaftlichen Entwicklung verändert, sind die Bewertungsgrößen $AT(M, D, t)$, $EK(M, D, t)$ und $C(M, D, t)$ grundsätzlich zeitabhängig. Der Zeitindex t wird im Folgenden ausdrücklich verwendet, wenn die Dynamik von $E(t)$, die Veränderung verfügbarer Alternativen oder die historische Verschiebung von Modellgeltung im Vordergrund steht. Wo es den Lesefluss stören würde, kann der Index t aus Gründen der Lesbarkeit weggelassen werden; inhaltlich ist er jedoch stets mitzudenken.

4. Kriterien moderner Modellbewertung im vorgeschlagenen Rahmen

Moderne wissenschaftliche Modellbewertung umfasst eine Reihe epistemischer, technischer und institutioneller Kriterien, die zusammen bestimmen, ob ein Modell innerhalb einer Domäne D erhalten bleibt, rekali­briert wird oder verschwindet. Der hier entwickelte Rahmen macht diese Kriterien explizit und ordnet sie den Bestandteilen der Nutzenfunktion $U(M, D, t)$ sowie dem epistemischen Ermög­lichungsraum $E(t)$ systematisch zu. Der folgende Katalog beansprucht nicht, alle disziplinspezifischen Bewertungsformen vollständig zu erfassen, sondern die zentralen epistemisch relevanten Bewertungsdimensionen moderner Modellpraxis sichtbar zu machen.

(1) Approximation und strukturelle Passung

Hierzu gehört die domänenspezifische Approximate Truth $AT(M, D, t)$, verstanden als graduelle Passung zwischen Modellstruktur und relevanten Vergleichsmerkmalen innerhalb einer Domäne. Dazu zählen empirische Fit-Qualität, Übereinstimmung mit robusten Datenstrukturen, strukturelle Ähnlichkeit, Stabilität der Modellvorhersagen und gegebenenfalls die korrekte Abbildung interventionell prüfbarer Zusammenhänge. AT bezeichnet damit die Passungsdimension eines Modells: Sie gibt an, wie gut ein Modell innerhalb einer Domäne zu den dort relevanten Merkmalen passt.

(2) Explanatorische Kraft (EK)

Die explanatorische Kraft eines Modells bestimmt, welche erklärende Ordnungsleistung ein Modell über bloße Passung hinaus erbringt. Sie umfasst die Fähigkeit, Phänomene verständlich zu machen, unterschiedliche Befunde zu vereinheitlichen, kausale oder interventionelle Abhängigkeiten sichtbar zu machen, gegenfaktische Einsichten zu ermöglichen und eine Domäne theoretisch zu strukturieren. $EK(M, D, t)$ ist daher nicht auf Prognosegenauigkeit, Fit oder strukturelle Ähnlichkeit reduzierbar. Zwei Modelle können eine ähnliche Approximate Truth besitzen, sich aber deutlich darin unterscheiden, wie viel sie erklären, ordnen oder theoretisch erschließen.

(3) Modellunsicherheit

Modellunsicherheit umfasst die Varianz, Sensitivität und Stabilität der Modellvorhersagen innerhalb einer Domäne. Hohe Unsicherheitsbreiten mindern den effektiven epistemischen Wert eines Modells, selbst wenn der mittlere Fit hoch ist. Sie beeinflusst sowohl $AT(M, D, t)$ als auch die epistemischen Kosten $C(M, D, t)$.

(4) Interpretierbarkeit und Transparenz

Modelle unterscheiden sich in ihrer strukturellen Zugänglichkeit, nachvollziehbaren Mechanismen, Testbarkeit und den Möglichkeiten der Diagnose von Fehlerquellen. Geringe Interpretierbarkeit erhöht die epistemischen Kosten und senkt die explanatorische Nutzbarkeit eines Modells.

(5) Technische und rechnerische Kosten

Dazu zählen Rechenaufwand, Energieverbrauch, numerische Komplexität, Datenbedarf und Implementierbarkeit. Diese technischen Kosten bilden einen zentralen Bestandteil von $C(M, D, t)$ und beeinflussen die Rationalität der Modellwahl in datenintensiven Wissenschaften erheblich.

(6) Institutionelle und infrastrukturelle Bedingungen

Modelle werden durch Normen, Software-Ökosysteme, Ausbildungsstrukturen, Datenformate und organisatorische Praktiken stabilisiert. Diese Faktoren sind im epistemischen Ermöglichungsraum $E(t)$ verortet und erklären, warum Modelle oft persistieren, selbst wenn sie in einzelnen Domänen suboptimal werden.

(7) Verfügbarkeit epistemischer Alternativen

Die Menge der im Ermöglichungsraum $E(t)$ verfügbaren Alternativmodelle bestimmt, ob kontextuelle Falsifikation zu Modellrevision, Domänenbegrenzung oder globaler Eliminierung führt. Modelle bleiben bestehen, solange sie in mindestens einer relevanten Domäne epistemisch optimal oder konkurrenzfähig sind. Entscheidend ist daher nicht nur die interne Leistung eines Modells, sondern seine relative Stellung gegenüber verfügbaren, prüfbar und stabilisierbaren Alternativen.

Dieser Kriterienkatalog präzisiert, welche Faktoren in der Nutzenfunktion $U(M, D, t)$ und im Ermöglichungsraum $E(t)$ operative Bedeutung besitzen. Zugleich ergibt sich daraus eine systematische Übersicht zentraler Bewertungsdimensionen, die moderne wissenschaftliche Modellpraxis prägen. Damit wird der integrierte Rahmen theoretisch anschlussfähig und empirisch interpretierbar, ohne einen vollständigen Katalog aller disziplinspezifischen Bewertungsformen zu beanspruchen.

5. Proposition: Zweistufige Falsifikationsstruktur

In der hier vorgeschlagenen Terminologie bezeichnet „Falsifikation“ nicht mehr eine binäre Wahrheitsentscheidung über M insgesamt, sondern eine systematisch rekonstruierbare Verschiebung der Nutzenwerte $U(M, D, t)$ über Domänen hinweg, die unter bestimmten Bedingungen zur vollständigen Aufgabe von M führt.

Für jedes Modell M gilt:

1. Solange mindestens eine relevante Subdomäne existiert, in der M gemäß $U(M, D, t)$ epistemisch optimal oder konkurrenzfähig ist, führt Falsifikation nicht zur globalen Eliminierung, sondern zur Einschränkung oder Revision des Geltungsbereichs:
 $D(M) \rightarrow D(M)'$, wobei $D(M)'$ die verbleibenden relevanten Domänen bezeichnet, in denen M epistemisch optimal oder konkurrenzfähig bleibt.
2. M ist genau dann global falsifiziert, wenn keine relevante Domäne D existiert, in der M gemäß $U(M, D, t)$ epistemisch optimal oder konkurrenzfähig bleibt.

Approximate Truth $AT(M, D, t)$, explanatorische Kraft $EK(M, D, t)$ und Modellkosten $C(M, D, t)$ bestimmen die Bewertung epistemischer Optimalität; $E(t)$ bestimmt die Menge verfügbarer und stabilisierbarer Alternativen.

Formal lässt sich der Übergang von kontextueller zu globaler Falsifikation als Verschiebung von relativer Konkurrenzfähigkeit zu stabiler Nicht-Konkurrenzfähigkeit ausdrücken. Dabei bezeichnet M^* jeweils das beste innerhalb der betreffenden Domäne verfügbare Alternativmodell. Kontextuelle Falsifikation liegt vor, solange mindestens eine relevante Domäne D existiert, in der M innerhalb des Toleranzbereichs ε konkurrenzfähig bleibt:

$$\exists D: U(M, D, t) + \varepsilon \geq U(M^*, D, t).$$

Globale Falsifikation liegt dagegen vor, wenn M in allen relevanten Domänen stabil unterhalb dieses Toleranzbereichs liegt:

$$\forall D: U(M, D, t) + \varepsilon < U(M^*, D, t).$$

Der Übergang ist damit nicht bloß ein Wechsel von Wahrheit zu Falschheit, sondern eine Veränderung relativer Modellgeltung gegenüber verfügbaren Alternativen.

Diese zweistufige Struktur macht eine in der Literatur zwar diskutierte, jedoch selten explizit modellierte Unterscheidung formal greifbar. Sie rekonstruiert präzise, wie Modelle in bestimmten Subdomänen erhalten bleiben können, obwohl sie in anderen scheitern, und ergänzt damit bestehende Ansätze von Popper und Lakatos durch eine klar strukturierte Darstellung domänenspezifischer Modellstabilität. Die zweistufige Falsifikationsstruktur beschreibt damit ausdrücklich die in der Wissenschaftspraxis häufig beobachtete Situation, dass Modelle in Subdomänen rekaliibriert werden, während ihre übergeordnete epistemische Rolle erhalten bleibt.

5.1 Begrenzung von Immunisierungswegen

Die theoretische Belastbarkeit der zweistufigen Falsifikationsstruktur hängt davon ab, dass drei systematisch verbundene Immunisierungswege ausgeschlossen werden. Erstens darf der Toleranzbereich ϵ nicht zu einer unbegrenzten Schutzklausel für geschwächte Modelle werden. Er bezeichnet nur jene Zone, innerhalb derer Konkurrenzfähigkeit nach den in einer Domäne anerkannten Fehler-, Unsicherheits- und Bewertungsstandards weiterhin plausibel bleibt. Zweitens darf die Relevanz einer Domäne nicht aus dem Verteidigungsinteresse des jeweils untersuchten Modells abgeleitet werden. Eine relevante Domäne muss durch bestehende wissenschaftliche Praxis, Zielgrößen, Prüfverfahren, Anwendungskontexte oder anerkannte Bewertungsordnungen unabhängig stabilisiert sein. Drittens darf Domänenzerlegung nicht als Ad-hoc-Rettung fungieren. Sie ist nur dann epistemisch legitim, wenn sie neue prüfbare Unterschiede sichtbar macht, Fehlerquellen diagnostisch trennt, Modellanwendungen präzisiert oder neue Erklärungs-, Prognose- oder Interventionsleistungen ermöglicht.

Diese drei Begrenzungen wirken zusammen. Eine Theorie kontextueller Falsifikation, die nur eine dieser Grenzen einzieht, bleibt offen für Immunisierung über die anderen: über ein zu weites ϵ , über künstlich erzeugte Restdomänen oder über bloß defensive Subdomänenbildung. Kontextuelle Falsifikation ist daher nur dann rational rekonstruierbar, wenn Toleranzbereiche, Domänenrelevanz und Domänenzerlegung jeweils durch externe disziplinäre Standards, bestehende Praxisformen und neue epistemische Anschlussleistungen begrenzt werden. So könnte etwa ein Modell, dessen Toleranzbereich ϵ disziplinär gut begrenzt ist, dennoch durch eine bloß defensive Subdomänenbildung erhalten werden, wenn diese nicht durch unabhängige Praxisstandards oder neue epistemische Anschlussleistungen gedeckt ist.

6. Wann Modelle verschwinden

Ein Modell verschwindet nicht bereits dann, wenn es in einzelnen Domänen Schwächen zeigt oder gegenüber einer Alternative geringfügig schlechter abschneidet. Globale Modellablösung setzt voraus, dass M unter den Bedingungen von $E(t)$ in keiner relevanten Domäne epistemisch optimal oder konkurrenzfähig bleibt. Entscheidend ist daher stabile relative Nicht-Konkurrenzfähigkeit gegenüber verfügbaren Alternativen. Diese Diagnose setzt voraus, dass weder ϵ künstlich erweitert noch Domänenrelevanz bloß aus dem Verteidigungsinteresse des Modells abgeleitet noch Domänenzerlegung als rein defensive Rettungsstrategie verwendet wird.

Ein Modell verschwindet typischerweise, wenn:

1. in allen relevanten Domänen mindestens ein alternatives Modell M^* dauerhaft einen Nutzensvorsprung jenseits des Toleranzbereichs ϵ aufweist:
 $U(M, D, t) + \epsilon < U(M^*, D, t)$,
2. die Modellkosten von M relativ zu seinen Alternativen nicht mehr durch Approximate Truth oder explanatorische Kraft kompensiert werden,
3. unter den Bedingungen von $E(t)$ keine aktuellen wissenschaftlichen Nutzungskontexte oder relevanten Domänen mehr bestehen, in denen M epistemisch konkurrenzfähig bleibt.

Beispiele:

- Äther → Maxwell (1865) und Einstein (1905),
- Phlogiston → moderne Chemie,
- Epizyklen → Kepler (1609, 1619) und Newton,
- Vier-Säfte-Lehre → moderne Medizin.

Diese historischen Modellwechsel werden in der Wissenschaftstheorie häufig als Paradigmenwechsel oder Forschungsprogramm-Transformationen interpretiert (Kuhn 1962; Lakatos 1970), finden jedoch in der hier vorgeschlagenen Nutzenstruktur eine präzisere formale Rekonstruktion.

Technologische, mathematische und institutionelle Innovationen verändern $E(t)$ und beschleunigen globale Falsifikation. Damit bestimmt $E(t)$ nicht nur, welche Modelle verschwinden, sondern auch, welche überhaupt als realistische Alternativen verfügbar sind.

7. Warum Modelle bleiben

7.1 Robustheit durch Modellfamilien

Viele Modelle existieren nicht als Einzelstrukturen, sondern als ganze Modellfamilien. Falsifikation betrifft daher meist Varianten oder spezifische Parametrisierungen, nicht die gesamte Modellklasse. Modellfamilien besitzen strukturelle Redundanzen, die es erlauben, Fehler in einigen Teilmodellen zu kompensieren, ohne den Gesamtansatz aufzugeben. Dies erklärt, warum Falsifikation oft lediglich zu lokaler Revision, interner Regewichtung oder Einschränkung des Geltungsbereichs führt, nicht zu einer Eliminierung der gesamten Klasse. Diese Beobachtung ist konsistent mit der modelltheoretischen Literatur, in der Modellfamilien als strukturierte Räume möglicher Varianten beschrieben werden (Weisberg 2013). Die hier eingeführte Domänenstruktur präzisiert diesen Ansatz, indem sie zeigt, wie Stabilität und Variation innerhalb einer Modellklasse formal voneinander abhängig sind.

7.2 Nutzungskontexte

Modelle bleiben aus unterschiedlichen Gründen erhalten. Einige dieser Gründe begründen epistemische Konkurrenzfähigkeit innerhalb relevanter Domänen, andere erklären lediglich historische, didaktische oder institutionelle Persistenz.

Modelle können weiterhin wissenschaftlich relevant bleiben, weil sie:

- in bestimmten Domänen erklärkräftig oder prognostisch brauchbar sind,
- rechnerisch effizient bleiben,
- technisch zuverlässig implementierbar sind,
- als stabile Bausteine innerhalb größerer Modellarchitekturen fungieren.

Daneben können Modelle auch didaktisch, historisch oder illustrativ fortbestehen. Diese Formen der Restnutzung erklären ihre kulturelle oder institutionelle Persistenz, begründen aber für sich genommen keine relevante Domäne im Sinne globaler epistemischer Konkurrenzfähigkeit. Technische Standardisierung kann diese Persistenz zusätzlich stabilisieren, begründet aber für sich genommen noch keine epistemische Konkurrenzfähigkeit.

Kontextuelle Falsifikation korrigiert Domänen, eliminiert aber das Modell nicht. Die Persistenz von Modellen in D_1 , trotz Falsifikationen in D_2 , ergibt sich damit aus der Nutzenstruktur $U(M, D_1, t)$ sowie aus der Stabilisierung institutioneller und technischer Kontexte im epistemischen Ermöglichungsraum $E(t)$. Die Existenz stabiler Nutzungskontexte ist eine Funktion des epistemischen Ermöglichungsraums $E(t)$: Institutionen, Datenformate, Softwarebibliotheken und Ausbildungsstrukturen sichern Modelle unabhängig von ihrer globalen Approximate Truth. Dadurch wird erklärbar, warum Modelle, deren Nutzenstruktur in einer Subdomäne D_2 geschwächt ist, in einer anderen Domäne D_1 weiterhin rational verwendet werden können. Dies entspricht der in vielen Wissenschaften beobachtbaren Praxis, Modelle als werkzeugartige Bausteine zu verwenden, deren Geltung kontextabhängig ist und nicht global beurteilt wird.

8. Fallstudie: Klimamodelle

8.1 Klimamodelle als Ensemble-Struktur

Klimamodelle sind keine Einzelmodelle, sondern komplexe Ensembles aus verschiedenen Modellvarianten und Szenarien (Oreskes et al. 1994):

$$\text{Ensemble} = \{M_1, M_2, \dots, M_n\}$$

Dieses Ensemble erlaubt die Kombination verschiedener physikalischer Annahmen, Parametrisierungen und Anfangsbedingungen.

Falsifikation betrifft daher selten das gesamte Ensemble, sondern spezifische Komponenten:

- Ozeanmodelle,
- atmosphärische Parametrisierung,
- Wolkenprozesse,
- biogeochemische Module.

Das Ensemble als Ganzes bleibt epistemisch stabil, selbst wenn einzelne Komponenten modifiziert oder ersetzt werden. Dies entspricht der etablierten Modellpraxis in der Klimawissenschaft, in der Ensembles ausdrücklich als epistemische Stabilitätsmechanismen fungieren (vgl. Oreskes et al. 1994; IPCC 2021). Die Domänenstruktur des Ensembles ist daher stabiler als die der Einzelmodelle.

8.2 Parametrisierung als epistemische Operation

Viele klimarelevante Prozesse können nicht aus grundlegenden physikalischen Gleichungen in voller Auflösung berechnet werden. Parametrisierungen dienen als epistemische Brücken zwischen physikalischer Theorie und numerischer Machbarkeit.

Falsifikation markiert typischerweise die Grenzen solcher Parametrisierungen. Dies führt zu lokalen Domänenanpassungen, in denen nicht notwendig das Gesamtmodell verworfen wird, sondern die Nutzenstruktur eines Teilmodells oder einer Parametrisierung innerhalb einer bestimmten Subdomäne sinkt:

$$U(M, D_2, t) \downarrow \rightarrow \text{Rev}(M, D_2) \text{ oder } D(M)' = D(M) \setminus D_2.$$

Parametrisierungen sind damit keine bloßen Näherungstechniken, sondern operationale Domänendefinitionen, deren Anpassung ein zentraler Mechanismus kontextueller Falsifikation ist. Der Gesamtansatz bleibt dabei erhalten, solange das Modell oder die Modellfamilie in anderen relevanten Domänen epistemisch konkurrenzfähig bleibt.

8.3 Bayesian Updating und adaptive Modellierung

In der Bewertung, Gewichtung und Kombination von Klimamodellen können bayesianische Verfahren eingesetzt werden, etwa im Kontext von Modellgewichtung, Ensemble-Auswertung oder probabilistischer Projektion:

$$\text{Posterior} \propto \text{Likelihood} \times \text{Prior}$$

Diese Formel dient hier nicht als Behauptung, dass Klimamodelle generell einem einheitlichen bayesianischen Aktualisierungsmechanismus folgen. Sie steht vielmehr idealtypisch für eine Klasse redistributiver Bewertungsverfahren, in denen empirische Abweichungen nicht notwendig zur Eliminierung eines Gesamtmodells führen, sondern die relative Gewichtung von Teilmodellen, Parametrisierungen oder Modellvarianten verändern können.

Eine schlechter prognostizierte Variable kann die Likelihood eines Teilmodells oder einer Parametrisierung reduzieren und damit seine domänenspezifische Approximate Truth $AT(M, D_2, t)$ schwächen. Dies führt jedoch nicht notwendig zur Eliminierung des gesamten Ensembles. Vielmehr können alternative Parametrisierungen, Modellvarianten oder Submodelle innerhalb des Modellportfolios stärker gewichtet werden, wenn sie in der betreffenden Domäne höhere epistemische Konkurrenzfähigkeit erreichen.

Zugleich ist in der realen Klimamodellierung Vorsicht geboten. Die Philosophie der Klimamodellierung hat wiederholt gezeigt, dass Ensemble-Übereinstimmung nicht unmittelbar in epistemische Sicherheit übersetzt werden darf. Modellgewichtungen sind umstritten, weil Klimamodelle häufig nicht statistisch unabhängig sind, gemeinsame Modellgenealogien besitzen, ähnliche Parametrisierungen teilen oder korrelierte Fehlerstrukturen aufweisen. Robustheit über mehrere Modelle hinweg ist daher nur dann stark belastbar, wenn sie nicht bloß dieselben Hintergrundannahmen, Datenstrukturen oder Fehlerquellen reproduziert, sondern durch hinreichend unabhängige Evidenzlinien, Modellvergleiche oder Prozessverständnisse gestützt wird. Genau deshalb ist Klimamodellierung ein besonders geeigneter Fall für die hier entwickelte Theorie: Sie zeigt, dass Modellgeltung nicht allein über Fit oder Ensemble Agreement bestimmt werden kann, sondern nur durch das Zusammenspiel von Approximate Truth, explanatorischer Kraft, Unsicherheit, Modellkosten, verfügbaren Alternativen und $E(t)$. Bayesian Updating wird hier daher nicht als universales Verfahren der Klimamodellbewertung eingeführt, sondern als formale Illustration dafür, wie Falsifikation innerhalb komplexer Modellportfolios redistributiv wirken kann. Für die hier entwickelte Theorie ist entscheidend, dass solche Verschiebungen als domänenspezifische Veränderung relativer Konkurrenzfähigkeit innerhalb eines Modellportfolios rekonstruiert werden können, unabhängig davon, ob die konkrete Disziplin bayesianische, frequentistische, heuristische oder ensemblebasierte Bewertungsverfahren verwendet.

8.4 Domänenstruktur

Die Funktionsbereiche von Klimamodellen lassen sich idealtypisch gliedern:

- D_1 : globale Temperaturtrends,
- D_2 : regionale Niederschlagsmuster,
- D_3 : Extremereignisse,
- D_4 : kurzfristige Klimavariabilität, etwa ENSO oder AMO.

Diese Domänen unterscheiden sich hinsichtlich Datenlage, räumlicher und zeitlicher Auflösung, Unsicherheitsbreite und prognostischer Stabilität. Globale Temperaturtrends gelten in vielen Bewertungszusammenhängen als robuster rekonstruierbar als regionale Niederschlagsmuster, Extremereignisse oder kurzfristige Variabilität. Daraus folgt jedoch keine einfache Hierarchie absoluter Modellgeltung. Vielmehr zeigt sich, dass die epistemische Tragfähigkeit von Klimamodellen domänen-, skalen- und szenarioabhängig bewertet werden muss. Gerade deshalb eignet sich die Klimamodellierung als Beispiel kontextueller Falsifikation: Abweichungen in einzelnen Subdomänen schwächen nicht notwendig die gesamte Modellfamilie, sondern führen häufig zu lokaler Revision, Regewichtung oder Domänenpräzisierung.

8.5 Miniaturisierte formale Illustration

Zur Verdeutlichung kann ein vereinfachtes Beispiel dienen. Aus Gründen der Übersichtlichkeit wird der Zeitindex t in dieser Miniatur weggelassen.

Seien zwei Modelle M und M^* verfügbar, die in zwei Domänen D_1 und D_2 verglichen werden.

$$AT(M, D_1) = 0.9, EK(M, D_1) = 0.8$$

$$AT(M^*, D_1) = 0.7, EK(M^*, D_1) = 0.8$$

$$AT(M, D_2) = 0.4, EK(M, D_2) = 0.5$$

$$AT(M^*, D_2) = 0.8, EK(M^*, D_2) = 0.9$$

Für moderate Kosten $C(M) \approx C(M^*)$ gilt:

- In D_1 bleibt M epistemisch optimal.
- In D_2 wird M^* optimal.
- Global folgt daraus keine Eliminierung von M , solange M in D_1 epistemisch optimal oder konkurrenzfähig bleibt.

Dies veranschaulicht den Grundsatz:

Kontextuelle Falsifikation schwächt die Nutzenstruktur eines Modells in einer bestimmten Domäne D_2 , ohne dass dadurch notwendig seine Geltung in D_1 oder seine Rolle innerhalb eines Modellportfolios zerstört wird. Diese Struktur ist charakteristisch für viele simulationsbasierte Wissenschaften, nicht nur für die Klimaforschung.

8.6 Konsequenzen für den epistemischen Status

Klimamodelle werden häufig nach dem Popper'schen Schema beurteilt:

„Eine fehlerhafte Vorhersage zeigt, dass das Modell falsch ist.“

Dies ist wissenschaftstheoretisch unzutreffend (Oreskes et al. 1994). Korrekt ist: Falsifikation wirkt domänenspezifisch und zeigt an, welche Subkomponenten weiterentwickelt werden müssen. Dies verdeutlicht, dass die Popper'sche Eliminationslogik für simulationsbasierte Modellarchitekturen unzureichend ist und durch eine domänenspezifische Nutzenstruktur ersetzt werden muss.

Die Fallstudie zu Klimamodellen dient als illustrative Anwendung einer modernen, daten- und simulationsintensiven Modellarchitektur. Sie zeigt exemplarisch, wie Ensemble-Strukturen, Parametrisierungen und adaptive Gewichtungen als Fälle kontextueller Falsifikation und domänenspezifischer Modellrevision rekonstruiert werden können. Die Struktur der Analyse ist grundsätzlich auf andere Modellarchitekturen übertragbar, etwa klassische physikalische Modelle, ökonomische Modelle, epidemiologische Modelle oder KI-Modellarchitekturen. Diese Übertragbarkeit ist jedoch als theoretische Anschlussfähigkeit zu verstehen, nicht als bereits abgeschlossene empirische Prüfung in allen genannten Bereichen.

9. Approximate Truth

Approximate Truth erklärt, warum Modelle trotz Fehlern epistemischen Fortschritt erzeugen können. Der Begriff wurde in der Wissenschaftstheorie besonders durch Niiniluoto (1987, 1998) und Oddie (1986) geprägt, wird hier jedoch ausdrücklich domänenspezifisch und rekonstruktiv verwendet. Approximate Truth bezeichnet in diesem Paper nicht den unmittelbaren Zugriff eines Modells auf ein Zielsystem an sich und auch nicht die gesamte epistemische Leistung eines Modells. Sie bezeichnet eine bestimmte Bewertungsdimension innerhalb von $U(M, D, t)$: die domänenspezifische Passung zwischen Modellstruktur und relevanten empirischen, strukturellen oder interventionell prüfbaren Vergleichsmerkmalen.

Dadurch wird Wahrheitsnähe nicht als globale Eigenschaft eines Modells verstanden, sondern als funktional rekonstruierbare Relation innerhalb bestimmter Geltungsbedingungen. Approximate Truth ist also enger zu fassen als epistemische Optimalität insgesamt. Ein Modell kann eine hohe Passung zu bestimmten Daten oder Strukturen besitzen und dennoch geringe explanatorische Kraft haben; umgekehrt kann ein theoretisch erklärkräftiges Modell epistemisch geschwächt sein, wenn seine Passung zu den relevanten Merkmalen einer Domäne unzureichend ist.

Formal lässt sich Approximate Truth als domänenspezifische Ähnlichkeits- und Passungsmetrik fassen:

$$AT(M, D, t) = \sum_i w_i \cdot \text{sim}(M, S_i, D, t).$$

Dabei bezeichnet S_i jeweils ein innerhalb der Domäne D als relevant bestimmtes Vergleichsmerkmal, etwa beobachtete Systemzustände, empirische Datenstrukturen, theoretisch ausgezeichnete Eigenschaften oder interventionell prüfbare Zielgrößen. Die Funktion $\text{sim}(M, S_i, D, t)$ beschreibt die Ähnlichkeit oder Passung zwischen Modell M und dem jeweiligen Vergleichsmerkmal S_i unter den zum Zeitpunkt t verfügbaren methodischen und empirischen Bedingungen. Die Gewichte w_i repräsentieren die relative Relevanz dieser Merkmale innerhalb der jeweiligen wissenschaftlichen Bewertungsordnung. In diesem formalen Rahmen ist Approximate Truth weder ein absolutes Wahrheitskriterium noch eine bloße Fit-Metrik, sondern ein domänenspezifisches Maß struktureller und empirischer Passung, das für die Modellwahl zentral ist.

Wichtig ist:

1. Approximate Truth ist immer domänenspezifisch.
2. Approximate Truth bezeichnet die Passungsdimension eines Modells, nicht seine gesamte epistemische Leistung.
3. Explanatorische Kraft bleibt analytisch eigenständig: Sie betrifft die erklärende, vereinheitlichende, kausale, gegenfaktische oder strukturierende Leistung eines Modells.
4. Modellkosten bleiben ebenfalls eigenständig: Sie betreffen die epistemischen, technischen und institutionellen Aufwände, Belastungen oder Abhängigkeiten eines Modells.
5. Falsifikation kann $AT(M, D_2, t)$ senken, $EK(M, D_2, t)$ schwächen oder $C(M, D_2, t)$ erhöhen; sie muss aber nicht automatisch die Geltung von M in anderen Domänen D_1 zerstören.
6. Globale Falsifikation kann nicht allein über Approximate Truth definiert werden. Entscheidend ist die gesamte Nutzenstruktur und die relative Konkurrenzfähigkeit gegenüber verfügbaren Alternativen:

$U(M, D, t) + \varepsilon < U(M^*, D, t)$ für alle relevanten Domänen D .

Die Kopplung von Approximate Truth an die Nutzenfunktion $U(M, D, t)$ zeigt, dass graduelle Wahrheitsnähe nicht isoliert wirkt, sondern im Wechselspiel mit explanatorischer Kraft und Modellkosten die epistemische Optimalität bestimmt. Approximate Truth wirkt innerhalb von Domänen und ergänzt die Theorie durch eine quantitative Dimension epistemischer Verbesserung.

Für die praktische Anwendung ist wichtig, dass sowohl die Ähnlichkeitsfunktion $\text{sim}(M, S_i, D, t)$ als auch die Gewichte w_i kontextabhängig bestimmt und je nach Disziplin unterschiedlich operationalisiert werden können. In vielen wissenschaftlichen Bereichen lässt sich $\text{sim}(M, S_i, D, t)$ durch etablierte Fit- und Fehlermetriken annähern, etwa durch Likelihood-Funktionen, Varianzmaße, Prognosefehler, Residuenanalysen oder Ähnlichkeitskennzahlen zwischen Modelltrajektorien und beobachteten Systemzuständen. In anderen Kontexten, insbesondere in theoretischen oder strukturell orientierten Wissenschaften, erfassen Ähnlichkeitsmetriken zusätzlich qualitative oder topologische Merkmale wie Symmetrieerhalt, Invarianzstrukturen oder die Reproduktion kausaler Abhängigkeitsmuster.

Die Gewichte w_i repräsentieren die relative Relevanz verschiedener innerhalb einer Domäne ausgewählter Vergleichsmerkmale. Sie können sich aus empirischen Bewertungspraktiken, disziplinären Standards oder modelltheoretischen Überlegungen ergeben. Damit ist Approximate Truth nicht auf eine einzelne Metrik festgelegt, sondern bildet eine anpassungsfähige Struktur, die sowohl numerische Genauigkeit als auch strukturelle Passung eines Modells berücksichtigt, ohne einen unmittelbaren Zugriff auf ein Zielsystem an sich vorauszusetzen.

10. Warum eine integrierte Theorie bisher fehlte

10.1 Historische Gründe

Vor dem digitalen Zeitalter wurden viele wissenschaftliche Modelle analytisch formuliert, sodass Domänenstrukturen weniger sichtbar waren. Erst numerische Simulationen und großskalige Datenmodelle machten kontextuelle Falsifikation zu einem zentralen Bestandteil wissenschaftlicher Praxis.

10.2 Logische Gründe

Poppers Ansatz war primär als methodologisches Abgrenzungs- und Prüfprinzip formuliert. Für die Analyse moderner Modellpraxis erscheint jedoch die daraus häufig abgeleitete binäre Eliminationsstruktur zu grob. Moderne Modelle sind häufig approximativ, domänenspezifisch und dynamisch rekalkulierbar. Ein rein binäres Verständnis von Falsifikation ist hierfür ungeeignet, weil Approximate Truth über Domänen hinweg unterschiedlich ausfallen kann und empirische Abweichungen nicht notwendig das Modell als Ganzes betreffen. Dies zeigt, dass klassische logische Falsifikationsstrukturen die operative Komplexität moderner Modellpraktiken nur begrenzt erfassen, solange sie graduelle Approximation, Domänenbindung und Relevanzgewichtungen nicht explizit abbilden.

10.3 Institutionelle Gründe

Wissenschaftliche Praxis ist eingebettet in:

- Rechenkapazitäten,
- Datennormen,
- Peer-Review-Strukturen,
- Forschungsförderung,
- technologische Infrastrukturen.

Diese Faktoren formen den epistemischen Ermöglichungsraum $E(t)$ und bestimmen damit, welche Modelle formulierbar und stabilisierbar sind. Poppers Ansatz modelliert diese Dimension nicht explizit. In der hier eingeführten Theorie ist $E(t)$ dagegen explizit als strukturierende Kraft institutioneller und technischer Rahmenbedingungen modelliert und bildet damit eine zentrale epistemische Variable.

11. Epistemischer Ermöglichungsraum

11.1 Dynamik des epistemischen Ermöglichungsraums

Der epistemische Ermöglichungsraum $E(t)$ verändert sich mit technischer, methodischer und institutioneller Entwicklung. Er ist dabei nicht nur ein Ermöglichungsraum, sondern zugleich ein Begrenzungsraum: Er bestimmt, welche Modelle formulierbar, prüfbar, vergleichbar und stabilisierbar sind, begrenzt aber ebenso, welche Alternativen überhaupt sichtbar, anschlussfähig oder institutionell akzeptabel werden. Ein Modell kann zu einem Zeitpunkt epistemisch optimal sein und zu einem späteren Zeitpunkt nicht mehr, ohne dass sich seine interne Struktur verändert hat. Entscheidend ist, dass sich $E(t)$ selbst verändert: neue Daten, Methoden, Rechenkapazitäten oder Standards verschieben die Menge verfügbarer Alternativen und damit die relative Modellgeltung. Zugleich wirken erfolgreiche Modelle auf $E(t)$ zurück, indem sie Datenformate, Recheninfrastrukturen, institutionelle Standards und Forschungspraktiken etablieren. Damit entsteht eine wechselseitige Dynamik: $E(t)$ ermöglicht und begrenzt Modellwahl, während stabilisierte Modelle den künftigen Ermöglichungsraum selbst mitprägen.

Wesentliche Einflussfaktoren sind:

1. Datenwachstum und verbesserte Messsysteme,
2. neue mathematische Methoden,
3. steigende Rechenleistung,
4. institutionelle Normierungsprozesse,
5. Entwicklung wissenschaftlicher Diskurse.

Diese Faktoren können als Veränderungsoperatoren $\Delta E(t)$ aufgefasst werden, die den epistemischen Ermöglichungsraum systematisch erweitern oder einschränken.

$E(t)$ strukturiert:

- die Menge der real verfügbaren Modelle,
- die Struktur ihrer Domänen,
- ihre Kosten $C(M, D, t)$,
- die verfügbaren Alternativen M^* .

Damit ist der epistemische Ermöglichungsraum der zentrale Rahmenparameter der Modellwahl. In wissenschaftshistorischer Perspektive erklärt der Begriff $E(t)$, warum Modellwechsel oft durch technologische und institutionelle Neuerungen ausgelöst werden, die erst neue Modellalternativen formalisierbar machen.

11.2 Struktur des epistemischen Ermöglichungsraums $E(t)$

Für eine präzisere Analyse lässt sich der epistemische Ermöglichungsraum $E(t)$ in drei funktionale Teilkomponenten gliedern, die zusammen bestimmen, welche Modelle formulierbar und stabilisierbar sind.

1. **$E_m(t)$** : methodisch-mathematische Bedingungen.

Dazu gehören verfügbare mathematische Verfahren, statistische Methoden, Modellierungsansätze und algorithmische Techniken. Sie bestimmen, welche Modelltypen überhaupt formuliert werden können und welche Approximationen zulässig sind.

2. **$E_t(t)$** : technische und datenbezogene Bedingungen.

Hierzu zählen Rechenleistung, Datenqualität, Softwareinfrastrukturen, numerische Tools und Simulationstechnologien. Diese Faktoren legen fest, welche Modelle praktisch implementierbar sind und mit welcher Auflösung, Stabilität oder Komplexität die Modellierung durchgeführt werden kann.

3. **$E_i(t)$** : institutionelle und organisatorische Bedingungen.

Diese Dimension umfasst wissenschaftliche Normen, Peer-Review-Verfahren, Förderlogiken, Ausbildungsstrukturen und etablierte Forschungspraktiken. Sie prägen, welche Modelle langfristig stabilisiert werden, welche Standards sich durchsetzen und welche Alternativen als wissenschaftlich akzeptabel gelten.

$E(t)$ ergibt sich aus dem Zusammenspiel dieser drei Dimensionen. Die Verfügbarkeit und Stabilisierung eines Modells ist damit nicht allein Ergebnis seiner Approximate Truth, explanatorischen Kraft oder Kosten, sondern auch der methodischen, technischen und institutionellen Rahmenbedingungen, die seine Formulierung, Implementierung und Weiterentwicklung ermöglichen oder begrenzen. Zugleich bestimmt $E(t)$, welche Alternativmodelle überhaupt als relevante Vergleichsmodelle auftreten können. Damit beeinflusst $E(t)$ nicht nur die absolute Bewertung eines Modells, sondern auch seine relative epistemische Konkurrenzfähigkeit gegenüber verfügbaren Alternativen.

12. Entscheidungslogik der Modellwahl

12.1 Struktur der Nutzenfunktion

Die Wahl zwischen zwei Modellen M und M^* lässt sich durch eine zeit- und domänenspezifische Nutzenfunktion rekonstruieren:

$$U(M, D, t) = \alpha \cdot AT(M, D, t) + \beta \cdot EK(M, D, t) - \gamma \cdot C(M, D, t).$$

Diese Funktion beschreibt nicht den absoluten Wert eines Modells, sondern rekonstruiert seine domänenspezifische epistemische Geltung unter endlichen Bedingungen. Sie ist keine Messformel erster Ordnung und kein algorithmisches Entscheidungsverfahren. Ihre Aufgabe besteht darin, vorhandene disziplinäre Bewertungspraktiken in einer gemeinsamen Vergleichsstruktur sichtbar zu machen. In diesem Sinn fungiert sie als rekonstruktive Bewertungsgrammatik zweiter Ordnung: Sie ordnet die Kriterien, nach denen Modelle in einer Domäne faktisch bewertet werden, ersetzt diese Kriterien aber nicht durch eine universelle Berechnungsformel.

Entscheidend ist daher nicht allein der isolierte Nutzenwert eines Modells, sondern seine Differenz zu verfügbaren Alternativen innerhalb derselben relevanten Domäne. Diese Differenz kann als Vergleichsstruktur formuliert werden:

$$\Delta U(M^*, M, D, t) = U(M^*, D, t) - U(M, D, t).$$

ΔU ist keine eigenständige zusätzliche Theorie, sondern eine rekonstruktive Vergleichsstruktur. Sie macht sichtbar, ob eine Alternative innerhalb einer Domäne nur geringfügig besser, deutlich überlegen oder epistemisch dominant wird. Damit verbindet ΔU die Nutzenfunktion mit der Frage, wann kontextuelle Falsifikation zu Revision, Domänenbegrenzung oder globaler Modellablösung führt.

Eine globale Bewertung kann durch eine aggregierte Funktion $U(M)$ erfolgen, die die domänenspezifischen Werte $U(M, D, t)$ über relevante Domänen hinweg gewichtet. Diese Aggregation bleibt jedoch sekundär gegenüber der domänenspezifischen Bewertung, weil Modellgeltung in diesem Rahmen primär innerhalb bestimmter Domänen bestimmt wird.

Dabei gilt:

- $AT(M, D, t)$: domänenspezifische Approximate Truth, also der Grad stabilisierter Passung zwischen Modellstruktur und relevanten empirischen, strukturellen oder interventionell prüfbaren Vergleichsmerkmalen innerhalb einer Domäne.
- $EK(M, D, t)$: explanatorische Kraft des Modells in Domäne D , also seine Fähigkeit, Phänomene zu erklären, Zusammenhänge sichtbar zu machen, kausale oder gegenfaktische Einsichten zu ermöglichen und eine Domäne theoretisch zu strukturieren.
- $C(M, D, t)$: Modellkosten, etwa Datenbedarf, Rechenaufwand, Interpretierbarkeit, institutionelle Infrastruktur, Implementierungsaufwand, Unsicherheit, Fehleranfälligkeit und Abhängigkeit von stabilisierenden Hintergrundbedingungen.

- α , β , γ : Gewichtungsfaktoren, die nicht als beliebige Setzungen einzelner Akteure zu verstehen sind, sondern als rekonstruierbare Gewichtungen disziplinärer Bewertungsordnungen innerhalb des epistemischen Ermöglichungsraums, insbesondere seiner methodischen und institutionellen Dimensionen $E_m(t)$ und $E_i(t)$.

Unterschiedliche Wissenschaften gewichten Approximate Truth, explanatorische Kraft und Modellkosten verschieden, etwa je nachdem, ob Prognosegenauigkeit, mechanistische Erklärung, rechnerische Effizienz, institutionelle Standardisierung oder praktische Anwendbarkeit im Vordergrund stehen. Die Parameter α , β und γ machen diese Bewertungsordnungen explizit, ohne sie frei zu erfinden.

Ein einfaches Beispiel macht diese Rekonstruktionslogik sichtbar. In einer prognoseorientierten Klimadomäne kann α besonders stark gewichtet werden, weil empirische Passung, Vorhersagestabilität und Unsicherheitskontrolle zentrale Bewertungsgrößen sind. In einer theoretisch-strukturellen Domäne kann β stärker ins Gewicht fallen, wenn es vor allem darum geht, Phänomene zu vereinheitlichen, kausale Abhängigkeiten sichtbar zu machen oder einen bisher ungeordneten Phänomenbereich theoretisch zu erschließen. In technischen Echtzeit- oder Anwendungskontexten kann γ besonders relevant werden, weil Rechenkosten, Datenbedarf, Robustheit, Implementierbarkeit und institutionelle Anschlussfähigkeit die tatsächliche Modellwahl stark begrenzen. α , β und γ sind daher keine frei erfundenen Parameter, sondern verdichten die in einer Domäne bereits wirksamen Zielgrößen, Prüfverfahren und Bewertungspraktiken zu einer expliziten Vergleichsstruktur.

Die Nutzenstruktur ist mit bestehenden Verfahren der Modellselektion kompatibel, ohne auf diese reduzierbar zu sein. Informationskriterien wie AIC und BIC operationalisieren in begrenzter Weise Abwägungen zwischen empirischer Passung und Modellkomplexität; Bayes-Faktoren und Bayesian Model Averaging bilden probabilistische Varianten von Modellvergleich und Modellgewichtung ab. Auch Severe Testing und Robustheitsanalysen leisten spezifische Formen wissenschaftlicher Modellbewertung, indem sie die Strenge von Prüfungen oder die Stabilität von Ergebnissen über Modellvarianten hinweg untersuchen. Die Mehrleistung der hier vorgeschlagenen Nutzenstruktur liegt nicht darin, solche Verfahren zu ersetzen. $U(M, D, t)$ rekonstruiert vielmehr die breitere Entscheidungslage, in der diese Verfahren wissenschaftlich wirksam werden: Welche Domäne wird bewertet? Welche Alternativen sind im epistemischen Ermöglichungsraum $E(t)$ tatsächlich verfügbar? Welche Modellkosten sind relevant? Welche explanatorische Leistung zählt innerhalb der betreffenden Domäne? Und wann führt lokale Unterlegenheit nur zu kontextueller Falsifikation, wann aber zu globaler Modellablösung? Die Nutzenfunktion ist daher kein Konkurrenzverfahren zur Modellselektion, sondern ein Rahmen zur Rekonstruktion der epistemischen Einbettung solcher Verfahren.

Viele Terme der Nutzenfunktion können in realen Wissenschaften durch empirische Proxy-Metriken oder etablierte quantifizierende Verfahren angenähert werden, etwa durch Fit-Metriken, Informationskriterien, Komplexitätsmaße, Likelihood-Strukturen, Prognosefehler oder Sensitivitätsanalysen. Diese Verfahren operationalisieren jeweils bestimmte Teilaspekte von $AT(M, D, t)$, $EK(M, D, t)$ oder $C(M, D, t)$ innerhalb einer konkreten Disziplin. Die Nutzenfunktion ist daher als idealtypischer Rekonstruktionsrahmen zu verstehen: Sie macht sichtbar, wie verschiedene Bewertungsdimensionen zusammenwirken, ohne zu behaupten, dass epistemische Geltung vollständig oder kontextfrei berechenbar wäre.

Die lineare Form der Funktion dient als heuristische Minimalform. Sie ist bewusst einfach gehalten, um die Grundstruktur der Abwägung sichtbar zu machen. In konkreten Anwendungen können jedoch nichtlineare, schwellenbasierte oder produktförmige Bewertungsformen angemessener sein. Dies gilt etwa dann, wenn explanatorische Kraft erst ab einem Mindestmaß an Approximate Truth epistemisch relevant wird, wenn Modellkosten ab bestimmten Schwellen dominieren oder wenn hohe Unsicherheit den Nutzen eines Modells nicht nur graduell, sondern strukturell begrenzt. Die lineare Darstellung ist daher nicht als Behauptung zu verstehen, dass wissenschaftliche Rationalität tatsächlich additiv berechenbar wäre, sondern als anschlussfähige Ausgangsform einer rekonstruktiven Modellbewertung.

Kontextuelle Falsifikation kann $AT(M, D_2, t)$ senken, $EK(M, D_2, t)$ schwächen oder $C(M, D_2, t)$ erhöhen, wodurch $U(M, D_2, t)$ sinkt, ohne notwendig $U(M, D_1, t)$ zu beeinträchtigen. In prozessualer Perspektive erscheint diese Veränderung als Verdichtung epistemischer Friktion: Die bisherige Modellstabilisierung erzeugt steigende Belastung, sinkende Erklärungskraft, wachsende Kosten oder zunehmende Abhängigkeit von Hilfsannahmen, ohne dass daraus automatisch globale Eliminierung folgt. Diese Friktion wird entscheidungsrelevant, wenn sie sich gegenüber verfügbaren Alternativen in ΔU als stabile relative Unterlegenheit zeigt. Falsifikation wirkt daher nicht automatisch global, sondern verändert zunächst die domänenspezifische Nutzenstruktur eines Modells.

Diese Struktur erklärt zudem, warum Modellfamilien robust bleiben können. Wenn $U(M_1, D_2, t)$ sinkt, können andere Modelle M_2 oder M_3 derselben Familie weiterhin höhere Werte in dieser Domäne aufweisen. Falsifikation wird so zu einem Mechanismus der internen Reallokation epistemischer Gewichte, nicht zu einer einfachen Eliminationslogik im Popper'schen Sinn.

12.2 Skalierung und empirische Verankerung der Nutzenfunktion

Damit die vorgeschlagene Nutzenfunktion $U(M, D, t)$ nicht bloß theoretisch bleibt, kann sie in vielen wissenschaftlichen Kontexten durch etablierte Bewertungsverfahren angenähert werden. Die Kalibrierung erfolgt relativ: Innerhalb einer Domäne D werden Modelle nicht an einem absoluten Maßstab gemessen, sondern im Verhältnis zu verfügbaren Alternativen und zu den in dieser Domäne anerkannten Gütekriterien verglichen.

Solche Gütekriterien können je nach Disziplin Fit-Metriken, Fehlerraten, Likelihood-Strukturen, Informationskriterien, Prognosequalität, Sensitivitätsanalysen, Komplexitätsmaße oder Interpretierbarkeitsanforderungen umfassen. Sie operationalisieren jedoch jeweils nur bestimmte Teilaspekte von $AT(M, D, t)$, $EK(M, D, t)$ oder $C(M, D, t)$. Die Nutzenfunktion bildet daher keine kontextfreie Messung epistemischer Geltung, sondern eine strukturierte Rekonstruktion disziplinärer Bewertungspraktiken.

Die empirische Verankerung liegt somit nicht darin, eine einheitliche Skala für alle Wissenschaften vorzugeben, sondern darin, vorhandene Bewertungsverfahren innerhalb einer Domäne systematisch aufeinander zu beziehen. Dadurch kann sichtbar werden, ob ein Modell stabilisiert, revidiert, domänenspezifisch eingeschränkt oder durch Alternativen ersetzt wird. Die Theorie besitzt insofern eine deskriptive und schwach normative Dimension: Sie beschreibt reale Bewertungspraktiken und macht zugleich explizit, unter welchen Bedingungen Modellentscheidungen rational nachvollziehbar sind.

12.3 Diagnoseprotokoll kontextueller und globaler Falsifikation

Aus der vorgeschlagenen Theorie ergibt sich ein Diagnoseprotokoll für wissenschaftliche Modellbewertung. Es dient nicht als mechanisches Entscheidungsverfahren, sondern als strukturierte Anwendungsschablone, mit der die impliziten Bewertungsdimensionen wissenschaftlicher Modellwahl explizit gemacht werden können. Das Protokoll soll klären, ob ein Modell stabil bleibt, lokal revidiert, auf bestimmte Domänen begrenzt, innerhalb einer Modellfamilie neu gewichtet oder global abgelöst wird.

1. **Modell M bestimmen**

Zunächst ist zu klären, was genau bewertet wird: ein einzelnes Modell, eine Modellfamilie, eine Parametrisierung, ein Submodell oder eine bestimmte Anwendung eines Modells. Diese Unterscheidung ist entscheidend, weil Falsifikation häufig nicht das gesamte Modell betrifft, sondern nur eine Variante oder einen spezifischen Einsatzbereich.

2. **Relevante Domäne D bestimmen**

Anschließend muss bestimmt werden, in welcher Domäne das Modell geprüft wird. Eine Domäne ist nicht bloß ein Thema oder Anwendungsfeld, sondern ein Bereich stabiler Geltungsbedingungen. Dazu gehören Zielgrößen, Datenlage, zulässige Idealisierungen, erwartete Genauigkeit, Prüfverfahren und praktische Anforderungen. Ohne Domänenbestimmung bleibt unklar, ob eine Abweichung ein lokales Problem oder ein globales Scheitern anzeigt.

3. **Epistemischen Ermöglichungsraum $E(t)$ rekonstruieren**

Im nächsten Schritt ist zu prüfen, welche methodischen, technischen und institutionellen Bedingungen zum Zeitpunkt t vorliegen. Dazu zählen verfügbare Daten, Messverfahren, Rechenkapazitäten, mathematische Methoden, Softwareinfrastrukturen, disziplinäre Standards und akzeptierte Bewertungspraktiken. $E(t)$ bestimmt nicht nur, welche Modelle möglich sind, sondern auch, welche Alternativen überhaupt sichtbar, vergleichbar und stabilisierbar werden.

4. **AT, EK und C getrennt erfassen**

Die Bewertung sollte Approximate Truth, explanatorische Kraft und Modellkosten zunächst getrennt rekonstruieren. AT betrifft die domänenspezifische Passung zwischen Modellstruktur und relevanten empirischen, strukturellen oder interventionell prüfbaren Vergleichsmerkmalen. EK betrifft demgegenüber die erklärende Leistung des Modells, also seine Fähigkeit, Zusammenhänge sichtbar zu machen, relevante Phänomene zu strukturieren und gegebenenfalls vereinheitlichende, kausale, gegenfaktische oder interventionelle Einsichten zu ermöglichen. C umfasst die epistemischen, technischen und institutionellen Kosten eines Modells, etwa Unsicherheit, Komplexität, Datenbedarf, Interpretierbarkeit, Rechenaufwand oder Infrastrukturabhängigkeit.

5. Alternativen M^* bestimmen

Falsifikation wird erst entscheidungsrelevant, wenn verfügbare Alternativen existieren. Deshalb ist zu klären, welche Alternativmodelle innerhalb derselben Domäne tatsächlich verfügbar, prüfbar, implementierbar und institutionell stabilisierbar sind. Die Bestimmung solcher Alternativen schließt an die Frage effizienter Suche unter endlichen Bedingungen an, da nicht alle logisch denkbaren Alternativen, sondern nur verfügbare, prüfbare und stabilisierbare Alternativen innerhalb von $E(t)$ entscheidungsrelevant sind (Rapp 2026d). Ein Modell kann trotz Schwächen stabil bleiben, wenn keine Alternative innerhalb von $E(t)$ eine höhere epistemische Konkurrenzfähigkeit erreicht.

6. Relative Differenz ΔU rekonstruieren

Nun wird nicht nur gefragt, ob M Schwächen besitzt, sondern wie groß der Unterschied zu M^* ist. ΔU zeigt, ob eine Alternative nur geringfügig besser, deutlich überlegen oder übermaßgebliche Prüf- und Anwendungskontexte hinweg stabil epistemisch dominant wird. Kleine Unterschiede können innerhalb des Toleranzbereichs ϵ liegen; deutliche und stabile Unterschiede können dagegen Revision, Domänenbegrenzung oder Modellablösung begründen.

7. Diagnose stellen

Aus den vorherigen Schritten ergibt sich die Diagnose. Bleibt M innerhalb D konkurrenzfähig, liegt Stabilisierung vor. Verliert M nur in einer Subdomäne D_2 an Leistung, liegt kontextuelle Falsifikation mit Revisionsbedarf vor. Bleibt M in D_1 tragfähig, aber nicht in D_2 , folgt Domänenbegrenzung. Wird innerhalb einer Modellfamilie ein Teilmodell geschwächt und ein anderes gestärkt, liegt interne Regewichtung vor. Verliert M in allen relevanten Domänen stabil seine Konkurrenzfähigkeit gegenüber verfügbaren Alternativen, liegt globale Modellablösung vor. Bevor eine dieser Diagnosen abschließend gestellt wird, ist zu prüfen, ob ϵ innerhalb disziplinärer Toleranzstandards bleibt, ob die zugrunde gelegten Domänen unabhängig von der Modellverteidigung stabilisiert sind und ob etwaige Domänenzerlegungen progressive Anschlussleistungen erbringen. Erst wenn diese drei Begrenzungen eingehalten sind, ist die Diagnose rational rekonstruierbar.

Dieses Diagnoseprotokoll übersetzt die formale Struktur des Papers in eine Anwendungspraxis. Es behandelt Falsifikation nicht isoliert als Beobachtungswiderspruch, sondern als diagnostischen Prozess, in dem Modellgeltung, Domäne, verfügbare Alternativen, Kosten und Ermöglichungsbedingungen gemeinsam rekonstruiert werden. Damit wird die Theorie nicht zu einem algorithmischen Entscheidungsverfahren, aber zu einem operativen Rahmen, der reale Modellentscheidungen transparenter und vergleichbarer macht.

13. Verhältnis zum wissenschaftlichen Realismus

Die integrierte Theorie steht zwischen strukturellem Realismus und instrumentellem Modellpluralismus. Diese Zwischenposition ist besonders für moderne Modellpraktiken relevant, da viele Wissenschaften weder vollständig realistisch noch vollständig instrumentalistisch operieren. Modelle besitzen in diesem Rahmen Wahrheitsnähe insofern, als sie innerhalb bestimmter Domänen stabile relationale, empirische oder interventionell relevante Strukturen erfolgreich erfassen und rekonstruierbar machen. Dieser Ansatz ist mit Formen des strukturellen Realismus kompatibel, weil er Wahrheitsnähe nicht im Gesamtmodell verortet, sondern in jenen stabilen relationalen Strukturen, die innerhalb spezifischer Domänen erhalten bleiben. Ihr Fortbestand hängt jedoch zugleich von praktischen Faktoren ab: Kosten, institutioneller Stabilität, verfügbaren Alternativmodellen und dem Wandel des Ermöglichungsraums $E(t)$.

Dieser Ansatz ergänzt den strukturellen Realismus, indem er explizit macht, dass Approximate Truth nicht die alleinige Determinante wissenschaftlicher Rationalität ist. Zugleich erweitert er instrumentellen Modellpluralismus, indem er formalisiert, wann Modelle trotz Teilfalsifikationen epistemisch optimal bleiben.

Der hier entwickelte Ansatz lässt sich innerhalb der Realismusdebatte als vermittelnde Position verorten. Er steht dem strukturellen Realismus, etwa in der Tradition von Worrall (1989) und Psillos (1999), insofern nahe, als Approximate Truth als Annäherung an stabile relationale Strukturen gedeutet wird. Zugleich verfolgt der Ansatz eine breitere Perspektive als der konstruktive Empirismus van Fraassens (1980), da er zeigt, dass die operative Rationalität wissenschaftlicher Modellpraxis nicht allein von empirischer Adäquatheit abhängt, sondern auch von institutionellen, technischen und explanatorischen Bedingungen, die im epistemischen Ermöglichungsraum $E(t)$ systematisch berücksichtigt werden.

In Abgrenzung zu da Costa und Frenchs Konzept partieller Wahrheiten (2003) integriert der vorliegende Rahmen Approximate Truth nicht nur als logische Struktur gradueller Wahrheit, sondern bettet sie in eine explizite Nutzenfunktion $U(M, D, t)$ ein, die explanatorische Kraft und Modellkosten umfasst. Dadurch wird wissenschaftliche Rationalität als Entscheidungsproblem innerhalb eines dynamischen Ermöglichungsraums rekonstruierbar. Der Ansatz positioniert sich somit zwischen strukturellem Realismus und instrumentellem Modellpluralismus, verbindet jedoch deren Einsichten zu einem integrativen Rahmen, der direkt an die empirische Modellpraxis moderner Wissenschaften anschließt.

Gerade in datenbasierten und simulationsintensiven Wissenschaften – etwa Klimaforschung, Ökonomie, Epidemiologie oder KI-Modellierung – liefert dieser Rahmen eine präzise Möglichkeit, Modellstabilität und Modellwechsel systematisch zu analysieren.

14. Falsifikation als domänenrelative Geltungsrevision

Falsifikation ist kein binärer Eliminationsmechanismus, sondern ein domänenspezifisches Werkzeug zur Präzisierung wissenschaftlicher Modelle. Die hier entwickelte Theorie fasst zentrale Einsichten aus Popper, Kuhn, Lakatos, da Costa/French und der modernen Modelltheorie in einem explizit formalisierten, einheitlichen Rahmen zusammen. Durch die Unterscheidung zwischen globaler und kontextueller Falsifikation wird sichtbar, dass empirische Abweichungen überwiegend zu einer Einschränkung des Geltungsbereichs führen, nicht zu einer vollständigen Aufgabe des Modells.

Die Integration von Approximate Truth, explanatorischer Kraft und Modellkosten in eine domänenspezifische Nutzenfunktion zeigt, dass Modellstabilität das Ergebnis relativer epistemischer Optimalität ist und nicht allein aus Wahrheitsnähe folgt. Modelle verschwinden erst dann, wenn sie unter den Bedingungen von $E(t)$ in keiner relevanten Domäne oder aktuellen wissenschaftlichen Nutzungspraxis mehr epistemisch konkurrenzfähig bleiben. Globale Modellablösung ist daher nicht die direkte Folge eines einzelnen empirischen Widerspruchs, sondern das Ergebnis stabiler relativer Nicht-Konkurrenzfähigkeit gegenüber verfügbaren Alternativen. Zugleich ist die Theorie kontextueller Falsifikation nur belastbar, wenn ϵ , Domänenrelevanz und Domänenzerlegung nicht als Immunisierungsinstrumente verwendet werden. Falsifikation wird damit nicht abgeschwächt, sondern präzisiert: Sie begrenzt, revidiert, gewichtet oder eliminiert Modellgeltung dort, wo die wissenschaftliche Praxis dafür tragfähige Gründe liefert.

Mit dem Begriff des epistemischen Ermöglichungsraums wird zudem erklärbar, warum technologische, institutionelle und methodische Veränderungen wissenschaftliche Modelllandschaften tiefgreifend beeinflussen. Modellwahl wird dadurch als dynamischer Prozess rekonstruierbar, in dem wissenschaftliche Rationalität nicht in einer singulären logischen Operation, sondern in einem strukturierten Zusammenspiel von Kontext, Approximation und institutioneller Stabilisierung besteht.

Der hier entwickelte Rahmen ist selbst als Modell zu verstehen. Seine primäre Domäne ist die Analyse wissenschaftlicher Modellgeltung unter endlichen Bedingungen, insbesondere in idealisierten, simulationsbasierten und modellpluralen Wissenschaften. Seine eigene Geltung ist daher ebenfalls domänenspezifisch: Er kann in bestimmten wissenschaftlichen Kontexten hohe rekonstruktive Leistung besitzen, ohne damit eine universelle Theorie wissenschaftlicher Rationalität zu sein. Seine Kosten liegen in Abstraktion, Formalisierungsaufwand und der Gefahr möglicher Scheinpräzision. Damit unterliegt auch die vorliegende Theorie den Bedingungen, die sie beschreibt: Sie kann kontextuell eingeschränkt, weiter präzisiert oder durch geeignetere Rahmen ergänzt werden.

Damit wird Falsifikation nicht abgeschafft, sondern präzisiert: Sie bleibt ein zentrales Rationalitätsinstrument wissenschaftlicher Praxis, wirkt jedoch nicht als einfacher binärer Ausschlussmechanismus, sondern als domänenrelative Anzeige begrenzter Modellgeltung. Ihre epistemische Funktion besteht darin, Modellgeltung zu begrenzen, zu revidieren, neu zu gewichten oder Modelle dort abzulösen, wo verfügbare Alternativen dauerhaft höhere epistemische Konkurrenzfähigkeit erreichen.

Lokale wissenschaftstheoretische Begriffsfassung dieses Papers

Die folgenden Begriffe stehen in einem Modulzusammenhang innerhalb des Epistemik-Programms. Friktion (Rapp 2026b), Revision (Rapp 2026c), Suche (Rapp 2026d) und Domänenstruktur (Rapp 2026e) sind in eigenständigen Arbeiten ausgearbeitet; das vorliegende Paper überträgt diese Architektur auf den Spezialfall wissenschaftlicher Falsifikation und ergänzt sie um die zweistufige Falsifikationsstruktur und den epistemischen Ermöglichungsraum. Die folgende Begriffsordnung dient der Stabilisierung zentraler Bedeutungen innerhalb dieses Textes. Sie erhebt keinen Anspruch auf Vollständigkeit und begründet keinen eigenständigen allgemeinen Begriffskanon. Begriffe, die hier nicht eigens aufgeführt sind, werden im Sinne des Epistemik-Basiskanons verwendet oder sind für den Funktionskern dieses Papers nicht zentral. Veränderungen, Präzisierungen oder Erweiterungen sind möglich, müssen jedoch ausdrücklich ausgewiesen, lokal begrenzt und begründet werden. Implizite Bedeutungsverschiebungen, stille Erweiterungen oder rückwirkende Umdeutungen sind ausgeschlossen.

Übernahme des Epistemik-Basiskanons

Dieses Paper übernimmt den im Epistemik-Basispaper definierten Begriffskanon als unveränderte Referenzbasis (Rapp 2026a). Die dort eingeführten Begriffe werden ohne Umdeutung und ohne implizite Verschiebung ihrer funktionalen Bedeutung verwendet. Dieses Paper führt keine abweichenden Definitionen der übernommenen Kanonbegriffe ein.

Lokale wissenschaftstheoretische Präzisierungen

Dieses Paper führt keine eigenständige allgemeine Kanon-Erweiterung ein. Es präzisiert einige übernommene Begriffe lokal für den Spezialfall wissenschaftlicher Falsifikation und moderner Modellbewertung. Diese Präzisierungen verändern den Epistemik-Basiskanon nicht, sondern übersetzen Teile seines Vokabulars in eine wissenschaftstheoretische Sprache für die Beschreibung von Modellgeltung, Domänenbegrenzung, Modellrevision und globaler Modellablösung.

Modellgeltung

Kurzdefinition: Domänenrelative epistemische Tragfähigkeit eines Modells unter bestimmten methodischen, technischen und institutionellen Bedingungen.

Funktion: Bezeichnet hier die Frage, in welchem Bereich ein Modell wissenschaftlich verwendbar, prüfbar und gegenüber Alternativen konkurrenzfähig bleibt.

Abgrenzung: Keine globale Wahrheit eines Modells; keine uneingeschränkte Gültigkeit; keine bloße faktische oder didaktische Nutzung.

Kontextuelle Falsifikation

Kurzdefinition: Domänenspezifische Schwächung oder Einschränkung von Modellgeltung, ohne dass das Modell als Ganzes eliminiert wird.

Funktion: Beschreibt Fälle, in denen ein Modell in einer Subdomäne revidiert oder begrenzt wird, während es in anderen relevanten Domänen weiterhin epistemisch tragfähig bleibt.

Abgrenzung: Keine vollständige Widerlegung des Modells; keine bloße Anomalie ohne Konsequenz; keine automatisch globale Modellablösung.

Globale Falsifikation

Kurzdefinition: Verlust epistemischer Konkurrenzfähigkeit in allen relevanten Domänen unter den Bedingungen des epistemischen Ermöglichungsraums $E(t)$.

Funktion: Markiert den Fall, in dem ein Modell gegenüber verfügbaren Alternativen keine relevante wissenschaftliche Tragfähigkeit mehr besitzt.

Abgrenzung: Keine bloße Teilfalsifikation; keine didaktische oder historische Nichtverwendung; keine einfache Folge eines einzelnen Beobachtungswiderspruchs.

Epistemischer Ermöglichungsraum $E(t)$

Kurzdefinition: Methodische, technische und institutionelle Bedingungsstruktur, innerhalb derer Modelle formulierbar, prüfbar, vergleichbar, stabilisierbar und Alternativen verfügbar werden.

Funktion: Erklärt, warum Modellgeltung historisch veränderlich ist und warum Alternativen erst unter bestimmten Bedingungen sichtbar, verfügbar und stabilisierbar werden.

Abgrenzung: Kein neutraler Hintergrundraum; keine bloße Sammlung externer Faktoren; keine mechanische Determination wissenschaftlicher Modellwahl.

Approximate Truth $AT(M,D,t)$

Kurzdefinition: Domänenspezifische Passung zwischen Modellstruktur, Beobachtungsdaten, Interventionspraxis und relevanten Vergleichsmerkmalen zum Zeitpunkt t .

Funktion: Dient der Rekonstruktion gradueller Wahrheitsnähe innerhalb bestimmter Geltungsbedingungen.

Abgrenzung: Keine globale Wahrheitseigenschaft eines Modells; kein unmittelbarer Zugriff auf ein Zielsystem an sich; keine bloße Fit-Metrik.

Explanatorische Kraft $EK(M,D,t)$

Kurzdefinition: Fähigkeit eines Modells, Phänomene innerhalb einer Domäne erklärend, strukturierend und gegenfaktisch aufschlussreich zu erfassen.

Funktion: Ergänzt Approximate Truth um die erklärende und strukturierende Leistung eines Modells.

Abgrenzung: Keine bloße Prognosegenauigkeit; keine rhetorische Plausibilität; keine domänenunabhängige Erklärungsmacht.

Modellkosten $C(M,D,t)$

Kurzdefinition: Epistemische, technische und institutionelle Aufwände oder Belastungen, die mit Formulierung, Anwendung, Stabilisierung oder Revision eines Modells verbunden sind.

Funktion: Macht sichtbar, dass Modellwahl nicht allein von Wahrheitsnähe oder Erklärungskraft abhängt, sondern auch von Unsicherheit, Komplexität, Datenbedarf, Rechenaufwand, Interpretierbarkeit und Infrastruktur.

Abgrenzung: Kein bloß monetärer Begriff; keine externe Zusatzbewertung; keine absolute Größe unabhängig von Domäne und Ermöglichungsraum.

Nutzenfunktion $U(M,D,t)$

Kurzdefinition: Rekonstruktive Bewertungsstruktur, die Approximate Truth, explanatorische Kraft und Modellkosten innerhalb einer Domäne zusammenführt.

Funktion: Dient der Darstellung domänenspezifischer epistemischer Optimalität oder Konkurrenzfähigkeit eines Modells.

Abgrenzung: Kein algorithmisches Entscheidungsverfahren; keine vollständige Berechnung wissenschaftlicher Rationalität; keine domänenunabhängige Rangordnung.

Relative Differenz $\Delta U(M^*,M,D,t)$

Kurzdefinition: Vergleichsgröße zwischen einem Modell M und einem verfügbaren Alternativmodell M^* innerhalb derselben relevanten Domäne.

Funktion: Macht sichtbar, ob eine Alternative nur geringfügig besser, deutlich überlegen oder stabil epistemisch dominant wird.

Abgrenzung: Keine eigenständige Theorie; keine exakte universelle Messgröße; keine Bewertung ohne Bezug auf $E(t)$.

Toleranzbereich ϵ

Kurzdefinition: Kontextabhängiger Bereich, innerhalb dessen ein Modell trotz geringfügig niedrigerem Nutzenwert weiterhin epistemisch konkurrenzfähig bleibt.

Funktion: Verhindert, dass minimale oder instabile Unterschiede zwischen Modellen sofort als globale Falsifikation gedeutet werden.

Abgrenzung: Keine willkürliche Schutzklausel; keine Immunisierung gegen Kritik; keine feste Größe unabhängig von Disziplin und Domäne.

Relevante Domäne

Kurzdefinition: Aktueller wissenschaftlicher Nutzungsbereich, in dem ein Modell erklärende, prognostische, interventionelle oder strukturierende Leistung erbringt.

Funktion: Unterscheidet epistemisch tragfähige Modellverwendung von bloß historischer, didaktischer oder illustrativer Restnutzung.

Abgrenzung: Kein beliebiger Anwendungskontext; keine rein kulturelle Persistenz; keine bloße Erwähnung eines Modells.

Status und Geltungsbereich

Die in diesem Paper vorgenommenen Präzisierungen stellen keine eigenständige Erweiterung des Epistemik-Basiskanons im starken Sinn dar. Sie dienen ausschließlich der lokalen wissenschaftstheoretischen Fassung zentraler Begriffe für den Spezialfall kontextueller und globaler Falsifikation wissenschaftlicher Modelle. Ihr Geltungsbereich ist auf die Modellbewertungsarchitektur dieses Papers begrenzt.

Es erfolgt keine stille Erweiterung, Umdeutung oder rückwirkende Modifikation des Epistemik-Basiskanons. Der Basiskanon bleibt in Bedeutung, Funktion und Abgrenzung unverändert bestehen.

Jede zukünftige Abweichung, weitergehende Präzisierung oder echte kanonische Erweiterung über diese lokale Begriffsfassung hinaus muss ausdrücklich ausgewiesen, lokal begrenzt und begründet werden. Implizite Bedeutungsverschiebungen oder informelle Kanonerweiterungen sind ausgeschlossen.

Literaturverzeichnis

- Cartwright, Nancy. 1983. *How the Laws of Physics Lie*. Oxford: Oxford University Press.
- Cartwright, Nancy. 1999. *The Dappled World: A Study of the Boundaries of Science*. Cambridge: Cambridge University Press.
- Da Costa, Newton C. A., und Steven French. 2003. *Science and Partial Truth: A Unitary Approach to Models and Scientific Reasoning*. Oxford: Oxford University Press.
- Duhem, Pierre. 1906. *La théorie physique: Son objet et sa structure*. Paris: Chevalier & Rivière.
- Einstein, Albert. 1905. „Zur Elektrodynamik bewegter Körper.“ *Annalen der Physik* 17: 891–921.
- Hacking, Ian. 1983. *Representing and Intervening*. Cambridge: Cambridge University Press.
- IPCC. 2021. *Climate Change 2021: The Physical Science Basis*. Cambridge: Cambridge University Press.
- Kepler, Johannes. 1609. *Astronomia Nova*. Heidelberg: Jonas Sabber.
- Kepler, Johannes. 1619. *Harmonices Mundi*. Linz: Johann Planck.
- Kitcher, Philip. 1981. „Explanatory Unification.“ *Philosophy of Science* 48 (4): 507–531.
- Kitcher, Philip. 1989. „Explanatory Unification and the Causal Structure of the World.“ In *Scientific Explanation*, herausgegeben von Philip Kitcher und Wesley C. Salmon, 410–505. Minneapolis: University of Minnesota Press.
- Kuhn, Thomas S. 1962. *The Structure of Scientific Revolutions*. Chicago: University of Chicago Press.
- Lakatos, Imre. 1970. „Falsification and the Methodology of Scientific Research Programmes.“ In *Criticism and the Growth of Knowledge*, herausgegeben von Imre Lakatos und Alan Musgrave, 91–196. Cambridge: Cambridge University Press.
- Maxwell, James Clerk. 1865. „A Dynamical Theory of the Electromagnetic Field.“ *Philosophical Transactions of the Royal Society of London* 155: 459–512.
- Morgan, Mary S., und Margaret Morrison. 1999. *Models as Mediators: Perspectives on Natural and Social Science*. Cambridge: Cambridge University Press.
- Niiniluoto, Ilkka. 1987. *Truthlikeness*. Dordrecht: Reidel.
- Niiniluoto, Ilkka. 1998. „Approximation in Science.“ *Poznan Studies in the Philosophy of the Sciences and the Humanities* 63: 97–109.
- Oddie, Graham. 1986. *Likeness to Truth*. Dordrecht: Reidel.
- Oreskes, Naomi, Kristin Shrader-Frechette, und Kenneth Belitz. 1994. „Verification, Validation, and Confirmation of Numerical Models in the Earth Sciences.“ *Science* 263 (5147): 641–646.
- Popper, Karl R. [1934] 1959. *Logik der Forschung*. Wien: Springer.
- Popper, Karl R. 1959. *The Logic of Scientific Discovery*. London: Hutchinson.

- Psillos, Stathis. 1999. *Scientific Realism: How Science Tracks Truth*. London: Routledge.
- Quine, W. V. O. 1951. "Two Dogmas of Empiricism." *The Philosophical Review* 60 (1): 20–43.
- Rapp, Stefan. 2026a. Epistemik: Modellmanagement unter endlichen Bedingungen. Zenodo. <https://doi.org/10.5281/zenodo.18441301>.
- Rapp, Stefan. 2026b. Friktion: Grenzsinal endlicher Tragfähigkeit in subjektiven, intersubjektiven und funktional-empirischen Stabilitätsräumen. Zenodo. <https://doi.org/10.5281/zenodo.18434648>.
- Rapp, Stefan. 2026c. Revision unter endlichen Bedingungen: Eine Theorie der Modelltransformation in der Epistemik. Zenodo. <https://doi.org/10.5281/zenodo.18935785>.
- Rapp, Stefan. 2026d. Effiziente Suche unter endlichen Bedingungen: Eine Dual-Mode-Architektur des Modellmanagements. Zenodo. <https://doi.org/10.5281/zenodo.18799355>.
- Rapp, Stefan. 2026e. Domänen, Grenzen und Übergangsfunktionen: Zur domänenrelativen Geltung, Migration und Kopplung von Modellen unter endlichen Bedingungen. Zenodo. <https://doi.org/10.5281/zenodo.19542403>.
- van Fraassen, Bas C. 1980. *The Scientific Image*. Oxford: Clarendon Press.
- Weisberg, Michael. 2013. *Simulation and Similarity: Using Models to Understand the World*. Oxford: Oxford University Press.
- Woodward, James. 2003. *Making Things Happen: A Theory of Causal Explanation*. Oxford: Oxford University Press.
- Worrall, John. 1989. „Structural Realism: The Best of Both Worlds?“ *Dialectica* 43 (1–2): 99–124.

Appendix: Didaktische Beispiele zur zweistufigen Falsifikationsstruktur

Die folgenden Beispiele dienen der didaktischen Veranschaulichung der im Paper entwickelten Unterscheidung zwischen kontextueller und globaler Falsifikation. Sie ersetzen keine fachhistorische Detailanalyse, sondern zeigen exemplarisch, wie Modellgeltung, Domänenrelevanz, Toleranzbereiche, Domänenzerlegung und globale Modellablösung im vorgeschlagenen Rahmen rekonstruiert werden können. Die Beispiele sind daher nicht als vollständige historische Rekonstruktionen zu verstehen, sondern als Anwendungsskizzen der begrifflichen Architektur dieses Papers.

A.1 Newtonsche Mechanik: kontextuelle Falsifikation ohne globale Eliminierung

Die Newtonsche Mechanik ist ein paradigmatischer Fall kontextueller Falsifikation. Ihre Geltung wurde durch relativistische Physik und moderne Gravitationstheorie nicht einfach global aufgehoben, sondern domänenspezifisch begrenzt. Bei sehr hohen Geschwindigkeiten, starken Gravitationsfeldern oder sehr präzisen astronomischen Messungen verliert sie gegenüber relativistischen Modellen epistemische Konkurrenzfähigkeit. In diesen Domänen sinkt ihre Approximate Truth, weil ihre Modellstruktur relevante Systemmerkmale nicht mehr hinreichend genau erfasst.

Gleichzeitig bleibt die Newtonsche Mechanik in vielen technischen, alltäglichen und ingenieurwissenschaftlichen Domänen weiterhin epistemisch tragfähig. Für Brückenbau, Maschinenbau, viele technische Anwendungen und klassische Näherungsberechnungen der Himmelsmechanik besitzt sie hohe explanatorische Kraft, niedrige Modellkosten und ausreichende Vorhersagegenauigkeit. Ihre schul- und universitätsdidaktische Nutzung kommt hinzu, begründet aber nicht für sich allein ihre epistemische Konkurrenzfähigkeit. In der Sprache dieses Papers bedeutet dies: Die Newtonsche Mechanik ist in bestimmten Subdomänen falsifiziert oder revisionsbedürftig, bleibt aber in anderen relevanten Domänen konkurrenzfähig.

Der Fall zeigt, warum Falsifikation nicht mit globaler Eliminierung gleichgesetzt werden darf. Die relativistische Physik ersetzt die Newtonsche Mechanik nicht in jedem Anwendungskontext, sondern begrenzt ihren Geltungsbereich. Die resultierende Struktur ist daher keine einfache Widerlegung, sondern eine Domänenpräzisierung: Das Modell bleibt erhalten, aber seine Geltung wird eingeschränkt.

Anti-Immunisierungsprüfung: Die Erhaltung der Newtonschen Mechanik ist keine bloße Schutzstrategie. Ihre verbleibenden Domänen sind unabhängig durch wissenschaftliche und technische Praxis stabilisiert. Ihre Fehlergrenzen sind bekannt, ihre Kosten gering, ihre Anwendungsbedingungen klar bestimmbar. Die Domänenbegrenzung erzeugt daher epistemische Anschlussleistung, statt ein geschwächtes Modell künstlich zu retten.

A.2 Phlogiston-Theorie: globale Falsifikation mit historischer Restnutzung

Die Phlogiston-Theorie eignet sich als Beispiel globaler Falsifikation. Sie verlor im Übergang zur modernen Chemie nicht nur in einer eng begrenzten Subdomäne an Leistung, sondern wurde in den relevanten wissenschaftlichen Kontexten der Erklärung von Verbrennung, Oxidation und chemischen Reaktionen durch leistungsfähigere Alternativen verdrängt. Ihre Modellkosten stiegen, ihre erklärende Kraft nahm ab, und verfügbare Alternativen boten stabil bessere Anschlussmöglichkeiten.

Im vorgeschlagenen Rahmen bedeutet dies: Die Phlogiston-Theorie blieb nicht in einer relevanten wissenschaftlichen Domäne epistemisch konkurrenzfähig. Ihre heutige Nutzung ist historisch, didaktisch oder wissenschaftsgeschichtlich, aber nicht mehr Teil aktueller chemischer Modellbewertung. Sie kann weiterhin erklärt, gelehrt oder als Beispiel wissenschaftlichen Wandels verwendet werden, doch diese Restnutzung begründet keine fortbestehende epistemische Konkurrenzzfähigkeit.

Der Fall macht die Unterscheidung zwischen historischer Persistenz und relevanter Domäne sichtbar. Ein Modell kann kulturell, didaktisch oder wissenschaftshistorisch präsent bleiben und dennoch global falsifiziert sein. Globale Falsifikation bedeutet daher nicht, dass ein Modell aus allen Diskursen verschwindet, sondern dass es in keiner aktuellen wissenschaftlichen Domäne mehr tragfähig gegenüber verfügbaren Alternativen bleibt.

Anti-Immunisierungsprüfung: Eine künstliche Restdomäne, etwa „Phlogiston als historische Denkfigur“, würde keine epistemische Konkurrenzzfähigkeit begründen. Sie ist keine aktuelle chemische Anwendungsdomäne, sondern eine didaktische oder historische Verwendungsform. Genau dadurch wird sichtbar, warum relevante Domänen nicht aus bloßer Restnutzung abgeleitet werden dürfen.

A.3 Klimamodelle: kontextuelle Falsifikation innerhalb von Modellfamilien

Klimamodelle zeigen, wie kontextuelle Falsifikation in komplexen Modellfamilien wirkt. Ein Klimamodell oder ein Ensemble von Klimamodellen kann in bestimmten Bereichen hohe Tragfähigkeit besitzen, etwa bei langfristigen globalen Temperaturtrends, während andere Subdomänen größere Unsicherheiten aufweisen, etwa regionale Niederschlagsmuster, Extremereignisse oder kurzfristige Klimavariabilität. Abweichungen in einer solchen Subdomäne falsifizieren nicht notwendig die gesamte Modellfamilie, sondern können einzelne Parametrisierungen, Teilmodelle oder Gewichtungen betreffen.

Im Rahmen dieses Papers heißt das: Falsifikation wirkt hier häufig redistributiv. Sie verschiebt epistemische Gewichte innerhalb eines Modellportfolios, schwächt bestimmte Komponenten und stärkt andere. Eine fehlerhafte regionale Prognose kann die Approximate Truth eines Teilmodells in einer spezifischen Domäne senken, ohne die gesamte Modellarchitektur global zu eliminieren. Die Folge ist dann lokale Revision, Regewichtung oder Domänenpräzisierung.

Dieser Fall zeigt besonders deutlich, warum moderne Modellpraxis nicht durch eine einfache Popper'sche Eliminationslogik erfasst werden kann. In simulationsbasierten Wissenschaften sind Modelle oft modular, parametrisiert und ensembleförmig organisiert. Falsifikation betrifft dann selten ein Gesamtmodell als einheitliche Struktur, sondern wirkt innerhalb eines gegliederten Modellraums.

Anti-Immunisierungsprüfung: Die Erhaltung eines Klimamodells oder Ensembles ist nur dann rational rekonstruierbar, wenn die verbleibenden Domänen unabhängig durch Daten, Prüfverfahren und Anwendungskontexte gestützt sind. Eine bloße Verschiebung problematischer Vorhersagen in immer kleinere Subdomänen wäre nicht ausreichend. Legitimität entsteht erst, wenn die Domänenpräzisierung diagnostisch, prognostisch oder erklärend neue Anschlussleistung erzeugt.

A.4 Ptolemäische Epizyklen: Grenzfall zwischen Anpassung und Ad-hoc-Stabilisierung

Ptolemäische Epizyklenmodelle eignen sich als Grenzfall, weil sie empirische Anpassungsfähigkeit und mögliche Ad-hoc-Stabilisierung zugleich sichtbar machen. Solche Modelle konnten beobachtete Planetenbewegungen über längere Zeit durch zusätzliche geometrische Konstruktionen beschreiben. In bestimmten historischen Ermöglichungsräumen waren sie daher nicht einfach irrational oder epistemisch leer. Ihre Stabilität beruhte auf mathematischer Anpassungsfähigkeit, institutioneller Einbettung und dem Fehlen konkurrierender, stabiler Alternativen mit höherer Gesamtleistung.

Mit der Entwicklung keplerischer und später newtonscher Modelle verschob sich jedoch der epistemische Ermöglichungsraum. Neue mathematische, empirische und theoretische Bedingungen machten Alternativen verfügbar, die nicht nur Beobachtungen reproduzierten, sondern auch eine tiefere strukturelle und dynamische Erklärung ermöglichten. Dadurch verlor die ptolemäische Modellklasse ihre relative Konkurrenzfähigkeit in relevanten astronomischen und physikalischen Domänen.

Dieser Fall zeigt die Bedeutung der Anti-Immunisierungslogik. Nicht jede Modellanpassung ist illegitim. Eine Anpassung ist epistemisch produktiv, wenn sie neue prüfbare Unterschiede sichtbar macht, Fehlerquellen trennt oder Erklärungskraft erhöht. Sie wird problematisch, wenn sie nur bekannte Gegenbefunde absorbiert, ohne neue Anschlussleistung zu erzeugen. Die Grenze zwischen legitimer Modellverfeinerung und Ad-hoc-Rettung liegt daher nicht in der bloßen Tatsache zusätzlicher Annahmen, sondern in deren epistemischer Produktivität.

Anti-Immunisierungsprüfung: Epizyklen werden dann problematisch, wenn zusätzliche Konstruktionen nur noch defensive Anpassungen leisten und keine neue explanatorische oder prognostische Struktur eröffnen. Ihre globale Ablösung ergibt sich nicht aus einer einzelnen Beobachtung, sondern aus stabiler relativer Unterlegenheit gegenüber Alternativen, die unter veränderten Bedingungen von $E(t)$ verfügbar und tragfähiger wurden.

A.5 Vier-Säfte-Lehre: globale Ablösung trotz kultureller Persistenz

Die Vier-Säfte-Lehre ist ein Beispiel dafür, dass ein Modell kulturell fortbestehen kann, ohne wissenschaftlich konkurrenzfähig zu bleiben. Historisch bot sie eine umfassende Ordnung medizinischer Beobachtungen, Krankheitsdeutungen und therapeutischer Praktiken. Innerhalb ihres damaligen Ermöglichungsraums konnte sie stabilisierend wirken, weil alternative physiologische, mikrobiologische oder biochemische Modelle noch nicht in heutiger Form verfügbar waren.

Unter modernen Bedingungen hat die Vier-Säfte-Lehre jedoch keine epistemische Konkurrenzfähigkeit in medizinischer Diagnostik, Therapie oder Krankheitsmodellierung. Ihre erklärende und interventionelle Leistung wird durch moderne Medizin, Physiologie, Pathologie, Mikrobiologie und evidenzbasierte Verfahren ersetzt. Eine heutige Berufung auf die Vier-Säfte-Lehre kann historisch, kulturgeschichtlich oder alternativmedizinisch vorkommen, begründet aber keine relevante wissenschaftliche Domäne im Sinne dieses Papers.

Der Fall verdeutlicht erneut den Unterschied zwischen Persistenz und Geltung. Modelle verschwinden nicht notwendig aus Sprache, Kultur oder Geschichte, wenn sie wissenschaftlich falsifiziert sind. Globale Falsifikation bedeutet vielmehr, dass ein Modell in den relevanten wissenschaftlichen Praxiszusammenhängen keine tragfähige Konkurrenzfähigkeit mehr besitzt.

Anti-Immunisierungsprüfung: Die Vier-Säfte-Lehre kann nicht dadurch epistemisch gerettet werden, dass man ihre heutige historische oder metaphorische Verwendung als relevante medizinische Domäne ausgibt. Eine solche Restdomäne wäre nicht unabhängig durch aktuelle medizinische Prüfverfahren, therapeutische Erfolge oder wissenschaftliche Bewertungsordnungen stabilisiert.

A.6 Übersicht der Beispiele

Fall	Problematische Domäne	Erhaltene Domäne	Diagnose im Rahmen dieses Papers
Newtonsche Mechanik	relativistische Geschwindigkeiten, starke Gravitation, hochpräzise Grenzfälle	technische, alltägliche und ingenieurwissenschaftliche Anwendungen sowie klassische Näherungsbereiche	kontextuelle Falsifikation mit Domänenbegrenzung
Phlogiston-Theorie	Verbrennung, Oxidation, chemische Reaktionslehre	keine aktuelle wissenschaftliche Domäne; historische und didaktische Restnutzung	globale Falsifikation
Klimamodelle	bestimmte regionale, kurzfristige oder hochunsichere Subdomänen	langfristige globale Trends und robuste ensemblegestützte Projektionen	kontextuelle Falsifikation innerhalb von Modellfamilien
Ptolemäische Epizyklen	astronomische Erklärung unter veränderten mathematischen und physikalischen Bedingungen	historische und didaktische Rekonstruktion	globale Ablösung; Grenzfall zwischen legitimer Anpassung und Ad-hoc-Stabilisierung
Vier-Säfte-Lehre	moderne medizinische Diagnostik und Therapie	historische, kulturgeschichtliche oder metaphorische Nutzung	globale Falsifikation trotz kultureller Persistenz

Die Beispiele zeigen, dass die zweistufige Falsifikationsstruktur nicht bloß eine formale Unterscheidung ist. Sie erlaubt es, verschiedene Fälle wissenschaftlicher Modellbewertung differenziert zu rekonstruieren: Modelle können lokal scheitern und dennoch rational erhalten bleiben; sie können historisch fortbestehen und dennoch global falsifiziert sein; sie können durch legitime Domänenpräzisierung verbessert oder durch bloße Ad-hoc-Zerlegung künstlich stabilisiert werden. Entscheidend ist jeweils, ob die verbleibende Modellgeltung durch relevante Domänen, disziplinäre Bewertungsstandards und neue epistemische Anschlussleistungen getragen wird.