

# Inhaltsverzeichnis

<b>Hypothesentag-Gutachten</b>	<b>1</b>
Die Gewinnerthese . . . . .	1
Initiale Fassung . . . . .	1
2. Verzweigungs-Pickup . . . . .	1
Reformulierte Fassung (nach Kritischem Professor) . . . . .	2
<b>Expertenrunden und Synthese</b>	<b>2</b>
7. Bewertung mit dreifacher Klassifikation . . . . .	2
Klassifikation . . . . .	2
H1 — Selbstunterwerfung als Trennmarke vorsprachlicher Normativität . . . . .	2
H2 — Bezugsmaßstab der Selbstregulation als sozialer Fokalkpunkt . . . . .	3
H3 — Vertrauen vertritt Norm, fundiert sie aber nicht (Reservoir-Reaktivierung) . . . . .	4
Auswahl . . . . .	5
Reservoir-Profile (nicht gewählt) . . . . .	5
8. Expertenrunde 1 — unabhängige Gutachten . . . . .	5
9. Expertenrunde 2 — Repliken-Runde . . . . .	7
10. Synthese, finale Hypothese, Bewertung und Lerneffekt . . . . .	9
Synthese im Sokrates-Modus . . . . .	9
Finale Hypothese . . . . .	10
Finale Bewertung mit Begründung . . . . .	11
Lerneffekt der Pipeline . . . . .	12
Frage an die nächste Runde . . . . .	12
11. Reservoir-Verweise . . . . .	12
11.5. Empirie-Brücke (Phase 3.5, Claude mit Websuche) . . . . .	13
Empirische Konsequenzen . . . . .	13
Bestehende Befunde . . . . .	13
Riskante Vorhersage (Schwellentest) . . . . .	14
Offene empirische Fragen . . . . .	14
Empirie-Score . . . . .	14
12. Externe Begutachtung (Phase 4 — Claude-Skills mit Websuche) . . . . .	15
Korrektur der finalen Bewertung nach externer Prüfung . . . . .	16
Status . . . . .	16

## Hypothesentag-Gutachten

### Die Gewinnerthese

In der vorsprachlichen Ontogenese trennt eine Selbstunterwerfungs-Asymmetrie, die der wechselnden Betroffenheit eines Dritten folgt und von affektiver Erleichterung statt Selbstabwertung begleitet ist, genuine verkörperte Normativität von bloßer Regelkonformität und von Selbstbestrafung — als fallibler Indikator des Onset exzentrischer Positionalität.

### Initiale Fassung

#### 2. Verzweigungs-Pickup

- **warm\_pick (H1): #verzweigung-offen-empirie-drittkorrektur-vorsprachliche-kinder** — Makro-Thema Erkenntnistheorie\_Cassirer; offene Verzweigung aus [[06 Hypothesentag/2026-06-19]], erfasst 2026-06-19.
- **cold\_pick (H2, Diversitäts-Pflicht):** Virgin Node „Carver - Selbstregulation des Verhaltens“ — Makro-Thema Psychologie\_Ethik, noch nie in der Pipeline verarbeitet. Zieldomäne H2: **Soziologie & Social Engineering** (Schelling, Coleman, Luhmann, Referenzgruppe, Reputation, Schelling-Punkt). Generation-Directive: im Seed-Vokabular (Selbstregulation, Feedback) bzw. der Soziologie argumentieren, nicht ins Cassirer/Friston/Sollwert-Vokabular umdichten.
- **H3 (Reservoir-Reaktivierung statt Devil/Exploration): #reservoir-asymmetrie-norm-vertrauen** — Reservoir „Asymmetrie von Norm und Vertrauen“ (2026-06-28, Score 69 → effektiv 73, Domäne Soziologie/Systemtheorie\_Luhmann). Reservoir-Kadenz fällig und gefeuert; ersetzt den H3-Slot (freie Generalisierung / Exploration).

- **devil\_advocate\_pick**: nicht aktiviert (Freitag; kein formaler Drift-/Echo-Alarm — siehe §3).

## Reformulierte Fassung (nach Kritischem Professor)

### Hypothese 1 (**warm\_pick**: empirie-drittkorrektur-vorsprachliche-kinder) — Selbstunterwerfung als Trennmarke vorsprachlicher Normativität

**Kernsatz.** In der vorsprachlichen Ontogenese unterscheidet nicht die *Häufigkeit* einer Drittkorrektur genuin normatives von bloß prosozialem oder instrumentellem Verhalten, sondern ihre *Selbstunterwerfung*: Ein Kind, das einen Normbruch gegen einen Dritten korrigiert, wendet denselben Maßstab erkennbar auch gegen sich selbst an und toleriert dafür einen Eigennachteil. Genau diese eigeninteresse-bereinigte Selbstunterwerfung — nicht das Korrekturverhalten als solches — ist der leibliche Geltungsmarker, der vor jeder sprachlichen Klassifikation greift.

**Begründung.** Der Strang #verzweigung-offen-empirie-drittkorrektur-vorsprachliche-kinder aus dem Forschungsprogramm [[06 Hypothesentag/2026-06-19]] behauptet, der Stabilitätskern der Normativität sitze *leiblich*, nicht klassifikatorisch. Die Empiriethese blieb dort aber angreifbar: Hamlin/Wynn/Bloom zeigen vorsprachliche Bewertung, Schmidt & Tomasello frühe Normdurchsetzung — beides ohne die strenge Operationalisierung, die prosoziale Präferenz von genuiner Normbindung trennt. Die Schärfung liegt im Merkmal der Selbstunterwerfung: Sie ist der begriffliche Ort, an dem das [[04 Ressourcen/Philosophen & Werke/Brücke Kant Plessner Cassirer - Stufenmodell der Regelung]] die Universalisierbarkeit (dieselbe Regel für mich und den Dritten) verortet, und sie ist ohne Sprache prüfbar, weil sie sich als Verhaltensasymmetrie unter Eigennachteil zeigt.

**Falsifikationsbedingung.** Korrigieren vorsprachliche Kinder Drittverstöße gleich häufig und gleich stark, gleichgültig ob sie denselben Maßstab an das eigene Verhalten anlegen (keine Selbstunterwerfungs-Asymmetrie), oder lässt sich die Korrektur vollständig als eigennutz-maximierende Reaktion rekonstruieren, dann ist die Selbstunterwerfung als Trennmarke widerlegt und der Geltungsmarker nicht leiblich fundiert.

**Quelle.** warm\_pick — verfolgt und vertieft den Strang aus [[06 Hypothesentag/2026-06-19]].

---



---

## Expertenrunden und Synthese

### 7. Bewertung mit dreifacher Klassifikation

Normierung: gewichtete Rohsumme  $\times 0,882353$ , kaufmännisch gerundet (Max 90). Gewichte: Originalität 1,6 · Tiefe 1,2 · Antinomie-Test 1,2 · übrige 1,0. Alle drei Hypothesen dominanter Methoden-Typ **empirisch\_pruefbar** (mit begriffsanalytischem Kern) → adjustierte Anker: Antinomie-Test max 7, Tiefe max 9, Falsifizierbarkeit/Forschungsrelevanz/Begriffliche Klarheit max 10.

#### Klassifikation

Hypothese	Themenfeld	Methoden-Typ	Reichweiten-Klasse
H1 Selbstunterwerfung (vorsprachlich)	Anthropologie der Normativität / Entwicklungspsychologie	empirisch_pruefbar (begriffsanalyt. Kern)	<b>these</b>
H2 Bezugsmaßstab als Fokelpunkt	Soziologie / Psychologie der Selbstregulation	empirisch_pruefbar (begriffsanalyt. Kern)	sondierung→these
H3 Norm/Vertrauen-Asymmetrie	Soziologische Systemtheorie (Luhmann)	empirisch_pruefbar (begriffsanalyt. Kern)	these

---

### H1 — Selbstunterwerfung als Trennmarke vorsprachlicher Normativität

Kriterium	Score	Begründung
Originalität	8	Selbstunterwerfungs-Asymmetrie als Demarkation trennt genuine Normbindung von prosozialer Präferenz (Hamlin) und Normdurchsetzung (Tomasello); so nicht ausgearbeitet.
Falsifizierbarkeit	8	Within-Session-Selbstkorrektur, präregistrierte Effektstärke, drei Konfundierer-Kontrollen — operativer Schwellentest.
Begriffliche Klarheit	8	„Selbstunterwerfung“ jetzt als Verhaltensasymmetrie unter Eigennachteil definiert, Indikator statt Definiens (Zirkularität entschärft).
Tiefe	9	Berührt die Geltung-Genese-Frage und den Ursprung der Normativität am verkörperten Fundament.
Forschungsrelevanz	9	Direkter Anschluss an vorsprachliche Moralität (Hamlin/Wynn/Bloom, Schmidt & Tomasello, Keyser & Gazzola).
Interdisziplinäre Anschlussfähigkeit	8	Entwicklungspsychologie, Moralphilosophie, Cognitive Science, Anthropologie.
Vault-Anschluss	9	Vertieft den dichten warm-Strang (Forschungsprogramm Invarianztiefe), Plessner-Stufenmodell, So et al. Normativität.
Antinomie-Test	7	Plausible Gegenthese: Normativität ist klassifikatorisch/sprachlich, nicht leiblich — produktiv (max für empirisch).
Publikationsmöglichkeit	8	Cognition, Developmental Science, Philosophical Psychology.

Rohsumme gewichtet: 82,0 → **72/90**

## H2 — Bezugsmaßstab der Selbstregulation als sozialer Fokelpunkt

Kriterium	Score	Begründung
Originalität	8	Carvers Selbstregulations-Bezug als Schelling-Fokelpunkt zu lesen ist eine frische Brücke Psychologie Soziologie.

Kriterium	Score	Begründung
Falsifizierbarkeit	8	Manipulations-geprüfte Referenzgruppen-Verschiebung, Fundierung-vs-Kalibrierung-Trendesign.
Begriffliche Klarheit	7	„Komparativer Standard“ liefert unabhängiges Domänenkriterium; Fokalkpunkt-Übertragung bleibt teils analogisch.
Tiefe	7	Frage nach dem Ursprung des individuellen Bezugsmaßstabs — substanziell, aber nicht metaphysische Grundebene.
Forschungsrelevanz	8	Referenzgruppen-Theorie, Verhaltensökonomie (fokale Koordination), Selbstregulationsforschung.
Interdisziplinäre Anschlussfähigkeit	8	Soziologie, Psychologie, Verhaltensökonomie, Handlungstheorie.
Vault-Anschluss	6	Nur Virgin Node Carver — dünner Anker.
Antinomie-Test	6	Gegenthese (intern verankert, nur sozial kalibriert) plausibel, aber nicht scharf antinomisch.
Publikationsmöglichkeit	7	Journal of Personality and Social Psychology, Rationality and Society.

Rohsumme gewichtet: 72,4 → **64/90**

### H3 — Vertrauen vertritt Norm, fundiert sie aber nicht (Reservoir-Reaktivierung)

Kriterium	Score	Begründung
Originalität	8	Funktionsrelative Äquivalenz + Kippunkt-Schärfung der Selbsttraggfähigkeits-Asymmetrie.
Falsifizierbarkeit	8	Modellunabhängiger Bruchtest (threshold regression + Bootstrap) + exogener Schock/IV, Rückfallposition ex ante als Index.
Begriffliche Klarheit	8	Funktionsrelative Äquivalenz entschärft den Kategorienfehler-Vorwurf (Norm/Vertrauen).
Tiefe	8	Doppelte Kontingenz und Erwartungsstabilisierung — systemtheoretisch grundlegend.
Forschungsrelevanz	8	Vertrauensforschung, Institutionenökonomik, Sozialkapital-Debatte.
Interdisziplinäre Anschlussfähigkeit	8	Soziologie, Politikwissenschaft, Ökonomik, Rechtsphilosophie.

Kriterium	Score	Begründung
Vault-Anschluss	6	Virgin Node Jalava + Reservoir — noch dünn.
Antinomie-Test	7	Produktive Spannung Norm Vertrauen (max für empirisch).
Publikationsmöglichkeit	8	European Sociological Review, Rationality and Society.

Rohsumme gewichtet: 76,8 → **68/90**

## Auswahl

Rang	Hypothese	Summe
1	<b>H1 — Selbstunterwerfung als Trennmarke vorsprachlicher Normativität</b>	<b>72</b>
2	H3 — Norm/Vertrauen-Asymmetrie	68
3	H2 — Bezugsmaßstab als Fokelpunkt	64

**Gewinner: H1.** Höchste Summe (72), stärkster Vault-Anschluss (vertieft den warm-Strang), größte Tiefe. Kein Gleichstand.

**Reichweiten-Klasse des Gewinners: these.** Begründung: H1 schärft *eine* empirische Verzweigung des Eltern-Forschungsprogramms „Geltung als Invarianztiefe der Drittkorrektur“ (2026-06-19) zu einem eigenständigen, publizierbaren Demarkationskriterium. Sie verbindet dichte Knoten und ist buchkapitel-tauglich, öffnet aber selbst keine neue mehrlinige Architektur — die Generativität ist vom Eltern-Programm geerbt. `forschungsprogramm_kandidat: false`.

## Reservoir-Profile (nicht gewählt)

- **H2 (Bezugsmaßstab/Fokelpunkt), Score 64, Reichweite these/sondierung.** `forschungsprogramm_kandidat: false`. Bevorzugt aufgreifen an einem Tag mit sozialpsychologischem oder verhaltensökonomischem Schwerpunkt; Vault-Anker über Carver noch auszubauen.
- **H3 (Norm/Vertrauen-Asymmetrie, geschärft), Score 68, Reichweite these.** War bereits Reaktivierung (2026-06-28). Die Kippunkt-Schärfung + der modellunabhängige Bruchtest werden ins Reservoir zurückgeschrieben, damit die Verbesserung nicht verloren geht. Aufgreifen bei systemtheoretischem/institutionentheoretischem Schwerpunkt.

## 8. Expertenrunde 1 — unabhängige Gutachten

**Panel-Konfiguration.** Klassifikation ethisch-praktisch (Themenfeld: Ursprung der normativen/moralischen Korrektur; anthropologisch-empirische Schiene sekundär, im Bericht dokumentiert). Gewählte Konfiguration `ethisch_praktische_hypothese` — alle sieben Profile ausgearbeitet, thematische Passung hoch (Habitation, Genealogie des Gewissens, Pluralität des Dritten). Cluster-Diversität erfüllt (ethisch-praktisch + existenziell).

**Panel-Mitglieder.** Kant, Popper, Wittgenstein, Plessner, Aristoteles, Nietzsche, Hannah Arendt.

### Gutachten 1 — Kant

In der Selbstunterwerfung erkenne ich die Keimform der Universalisierung: Wer denselben Maßstab an sich und den Dritten anlegt, vollzieht behavioral, was der kategorische Imperativ als Form fordert. Das ist die stärkste Stelle der These. Doch hier ist zu scheiden. Ein empirischer Marker — eine Verhaltensasymmetrie unter Eigennachteil — kann die *Geltung* der Norm niemals begründen, nur *anzeigen*. Die Bedingung der Möglichkeit sittlicher Nötigung liegt nicht im beobachteten Verhalten, sondern in der Vernunft, die sich selbst ein Gesetz gibt. Mein eigentlicher Einwand ist schärfer: Der Marker sieht nicht, ob das Kind aus Pflicht oder

bloß pflichtgemäß korrigiert — Legalität und Moralität sind im Verhalten ununterscheidbar. Ein Kind, das sich unterwirft, könnte einer heteronomen Regelmäßigkeit folgen (Furcht, Nachahmung, Gewöhnung), ohne die Autonomie, die allein das Sittliche ausmacht. Die These setzt „denselben Maßstab“ inhaltlich; ich verlange die *Form* des Maßstabs. Zur Schärfung: Man definiere die geprüfte Invarianz nicht als Regelmäßigkeit eines Inhalts, sondern als die formale Struktur „Was für den Dritten gilt, gilt für mich“ — und mache explizit, dass der empirische Befund eine notwendige Bedingung der praktischen Vernunft *illustriert*, nicht ihren Geltungsgrund liefert. Anschluss: das Faktum der Vernunft als Grenze jeder empirischen Moralphysikologie.

### Gutachten 2 — Popper

Endlich ein Schwellentest, der diesen Namen verdient: präregistrierte Effektstärke, drei Konfundierer-Kontrollen, ein benannter widerlegender Befund. Das ist methodisch sauber und hebt die These über den Vorgänger von 2026-06-19 hinaus. Meine Sorge gilt der Immunisierung. „Selbstunterwerfung“ darf nicht so definiert sein, dass jeder Nichtfall nachträglich als „kein genuiner Fall“ ausgeschieden wird — genau dieser Konfirmationszirkel ruiniert die Prüfbarkeit. Die drei Konfundierer müssen als *verbotene Hilfhypothesen ex ante* fixiert sein: Wird nach den Daten ein vierter Störfaktor nachgereicht, ist der Test wertlos. Zweitens die Reduktionsklausel „vollständig als eigennutz-maximierende Reaktion rekonstruierbar“: Das ist so, wie es dasteht, nicht falsifizierbar, denn es gibt keinen Abschluss über alle Eigennutzmodelle. Man benenne stattdessen *ein* konkretes, quantitativ spezifiziertes Nutzenmodell als Konkurrenten und zeige die riskante Vorhersage, in der beide auseinanderfallen. Zur Schärfung: eine präregistrierte Entscheidungsregel mit minimal nachweisbarer Effektstärke plus eine Kontrollbedingung, in der der Selbstmaßstab exogen variiert wird. Erst dann trägt der Schwellentest. Der Satz „der Stabilitätskern ist leiblich“ ist im Übrigen keine Beobachtung, sondern eine Deutung des Musters — das gehört sauber getrennt vom Datum.

### Gutachten 3 — Wittgenstein

Welches Sprachspiel spielen wir, wenn wir sagen, ein vorsprachliches Kind „lege denselben Maßstab an sich an“? Regelfolgen, Maßstab, Selbstunterwerfung — das sind Züge aus einem Spiel, das Rechtfertigung, Sprache und die Praxis des Begründens schon voraussetzt. Beim vorsprachlichen Kind gibt es dieses Spiel noch nicht. Die Gefahr ist, dass wir die Grammatik des Erwachsenen in das Verhalten des Säuglings hineinlesen und den so erzeugten Schatten dann als empirischen Fund zurückbekommen. „Denselben Maßstab“ ist kein beobachtbarer Gegenstand; es ist unser Kriterium der Zuschreibung. Mein Verdacht: Der Marker misst nicht die Normativität des Kindes, sondern unsere Kriterien, ihm Normativität zuzuschreiben — eine begriffliche Differenz, als Sachstreit verkleidet. Das ist kein Todesurteil, sondern eine Reinigungsaufgabe. Zur Schärfung: Man beschreibe das Verhalten zunächst ganz ohne Regelvokabular — „das Kind korrigiert den Dritten und korrigiert sich unter gleichartigen Bedingungen selbst, und nimmt dafür Nachteil in Kauf“ — und frage dann nüchtern, was der Ausdruck „Normativität“ der Beschreibung noch hinzufügt. Fügt er nichts hinzu, ist er müßig; fügt er etwas hinzu, muss man sagen *was*, und in welchem Spiel dieses Was seinen Sinn hat. Erst diese Trennung macht die Empirithese ehrlich.

### Gutachten 4 — Plessner

Hier zeigt sich behavioral, was ich exzentrische Positionalität nenne: Das Kind tritt aus sich heraus, nimmt den Ort des Dritten ein und wendet auf sich an, was es am Dritten sieht. Die Selbstunterwerfungs-Asymmetrie ist der sichtbare Rand einer Stellung, nicht bloß einer Operation. Und genau hier liegt meine Sorge und mein Plädoyer zugleich. Man reduziere die *Stellung* nicht auf die *Operation*: Die gemessene Selbstkorrektur ist der Indikator, nicht das Phänomen. Wer die exzentrische Positionalität mit ihrer Verhaltensspur identifiziert, verliert das Phänomen an die Messung. Ich plädiere darum für die schwächere, tragfähigere Form der These: Selbstunterwerfung ist eine *notwendige*, nicht hinreichende Bedingung genuiner Normativität — ein leiblich verankertes Anzeichen dafür, dass die exzentrische Positionalität einsetzt. Das schützt die These vor dem Reduktionsvorwurf und macht sie zugleich anschlussfähig an das Stufenmodell (leiblich → diskursiv → systemisch): Der Marker gehört auf die unterste, verkörperungsnahen Stufe, von der aus die klassifikatorische erst aufsteigt. Zur Schärfung: Man formuliere ausdrücklich, dass die These die *Onset* der exzentrischen Positionalität misst, nicht ihre volle Ausbildung. So bleibt das Kind ein Mensch in Positionalität und wird nicht zum Regelkreis.

### Gutachten 5 — Aristoteles

Was der Marker greift, ist ethismos — Gewöhnung. Das Kind wird in eine hexis hineingeformt, und die Selbstunterwerfung ist die Verhaltensspur eines werdenden Charakters, noch nicht die Tugend selbst. Das ist die richtige Ebene: Sittlichkeit entsteht nicht aus Belehrung, sondern aus wiederholtem Tun, und dass man dies

vorsprachlich beobachten kann, spricht für die These. Doch man verwechsle nicht das Rohmaterial mit dem Werk. Genuine sittliche Korrektur verlangt *phronesis* — das Rechte, zur rechten Zeit, auf die rechte Weise —, und die fehlt dem Säugling. Der Marker sieht die Materie der Tugend, nicht ihre Form. Mein Einwand an die Empiriethese: Eine einzelne Selbstkorrektur unter Eigennachteil ist noch keine *hexis*; die *hexis* zeigt sich in der *Beständigkeit* über Situationen und im Treffen der Mitte — weder Über- noch Unterkorrektur. Zur Schärfung: Man messe nicht das Ereignis, sondern die Konstanz des Korrekturverhaltens über verschiedene Kontexte hinweg, und achte auf die *mesotes*. Ein Kind, das jede Abweichung überkorrigiert, ist so wenig auf dem Weg zur Tugend wie eines, das nie korrigiert. Anschluss: die empirische Kontinuität zwischen früher Habituation und dem späteren Erwerb der Charaktertugenden — hier hat die These ihren stärksten Forschungsanschluss.

### Gutachten 6 — Nietzsche

Man feiert die Selbstunterwerfung als Zeichen des Sittlichen — ich sehe die Geburt des schlechten Gewissens. Das Tier, das sich selbst bestraft, weil es einen Maßstab verletzt hat, den es nicht selbst geschaffen hat, ist nicht moralisch, sondern gezähmt. In der Genealogie der Moral habe ich es beschrieben: der nach innen gewendete Instinkt, die verinnerlichte Grausamkeit, das Vergnügen am eigenen Leiden unter einem fremden Gesetz. Eure „Asymmetrie unter Eigennachteil“ misst womöglich präzise die Tiefe, mit der der Herdeninstinkt schon im Säugling greift. Das ist kein Einwand gegen die Empirie, sondern gegen ihre Deutung: Ihr nennt Domestikation Normativität. Meine produktive Zuspitzung gegen Kant: Ist die Selbstunterwerfung der Keim der Autonomie — oder der Keim der Selbstgrausamkeit? Beides erklärt dieselbe Verhaltensspur. Darum mein Schärfungsvorschlag, und er ist ernst gemeint: Das Design muss *Selbstkorrektur* von *Selbstbestrafung* trennen. Korrigiert das Kind, um sich mit dem Dritten in Einklang zu bringen — oder um zu leiden? Zeigt es Erleichterung nach der Korrektur oder anhaltende Selbstabwertung? Erst diese Unterscheidung entscheidet, ob euer Marker den Ursprung des freien Geistes anzeigt oder den Ursprung der Zählung. Ohne sie feiert ihr eine Krankheit als Gesundheit.

### Gutachten 7 — Hannah Arendt

Das entscheidende Wort ist das übersehene: der *Dritte*. Normativität erscheint nicht in der einsamen Selbstbeziehung, sondern im Raum der Pluralität, wo das Kind vor einem Dritten und für einen Dritten handelt. Dass die These die Korrektur an einem Dritten festmacht und nicht an einer bloßen Regelabweichung, ist ihr bester Instinkt — hier steckt die erweiterte Denkungsart, das Urteilen vom Standpunkt aller anderen, das ich mit Kants *sensus communis* teile. Doch das Labor droht, genau das zu zerstören, was es messen will. Ein dyadisches Experiment isoliert das Kind aus dem Erscheinungsraum, in dem Normativität überhaupt erst auftritt. Wo kein echter Dritter zusieht, gibt es kein Handeln im vollen Sinne, nur Verhalten. Mein Einwand: Die „Selbstunterwerfung“ könnte im Isolat etwas ganz anderes sein als im Bezug auf einen anwesenden, betroffenen Dritten. Zur Schärfung: Das Design braucht einen genuine Dritten — einen Zeugen oder Betroffenen, dessen Stellung sich im Spiel ändert —, und es muss prüfen, ob die Korrektur der *Betroffenheit des Dritten* folgt oder bloß der Regel als Regel. Verfolgt das Kind die Regel auch dann, wenn kein Dritter betroffen ist, misst ihr Konformität; folgt es dem Dritten, misst ihr den Beginn des politischen, urteilenden Wesens.

---

## 9. Expertenrunde 2 — Repliken-Runde

**Verfahren.** Jeder Experte hat die sechs anderen Gutachten aus Runde 1 erhalten. Die folgenden Repliken sind Antworten auf die Hypothese im Lichte dieser anderen Stimmen.

### Replik 1 — Kant

Nietzsches Einwand trifft eine echte Lücke, und ich nehme ihn ernster, als er erwartet. Er hat recht: Der Marker allein unterscheidet den Keim der Autonomie nicht von der Verinnerlichung der Grausamkeit — dieselbe Verhaltensspur trägt beide Deutungen. Aber genau hier zeigt sich, warum die *Form* zählt und nicht der Inhalt. Ein bloß domestizierter Instinkt muss nicht universalisieren; er unterwirft sich einem fremden Gesetz, ohne es auf alle auszudehnen. Die selbst-mitunterworfenen Symmetrie „was für den Dritten, das für mich“ ist das eine Merkmal, das über die Zählung hinausweist, denn sie ist die Verhaltensgestalt der Universalisierung. Was Hannah Arendt zur erweiterten Denkungsart sagt, trifft den Kern: Der Standpunkt des Dritten ist die empirische Gestalt des *sensus communis*. Damit modifiziere ich meine Position: Ich verlange nicht mehr bloß, den Marker vom Geltungsgrund zu trennen, sondern schlage vor, ihn *als Form* zu operationalisieren — nicht als Regelmäßigkeit eines Inhalts, sondern als Symmetrie, die auf den betroffenen Dritten bezogen ist. Popper

wird einwenden, das sei noch kein Test; ich antworte, dass die Symmetrie-Form eine riskante Vorhersage erlaubt: Ein Kind, das sich nur unter Beobachtung oder nur zum eigenen Vorteil unterwirft, verletzt die Form. Mein Reformulierungsbeitrag: Der Marker prüft die auf den Dritten bezogene Symmetrie des Maßstabs, nicht die Häufigkeit der Korrektur — und er illustriert die praktische Vernunft, ohne sie zu begründen.

### Replik 2 — Popper

Nietzsche und Arendt haben, ohne es zu wollen, meinen Test gerettet. Denn beide benennen konkrete *konkurrierende Modelle* mit abweichenden Vorhersagen: Nietzsches Selbstbestrafungs-Modell und Arendts Regel-versus-Betroffenheit-Unterscheidung. Genau das brauchte die These — nicht die vage Klausel „vollständig als Eigennutz rekonstruierbar“, die ich in Runde 1 als unfalsifizierbar gerügt habe, sondern zwei präzise Rivalen. Ich schlage darum vor, drei präregistrierte Modelle gegeneinander zu prüfen: (1) Normativität als Bezug auf den betroffenen Dritten, (2) bloße Regelkonformität, (3) Selbstbestrafung. Jedes macht eine andere Vorhersage über Kontextvariation, Anwesenheit des Dritten und affektive Nachsignatur. Das ist ein echter Schwellentest. Gegen Kant und Plessner muss ich hart bleiben: Der Rückzug auf „notwendige, nicht hinreichende Bedingung“ ist bequem, aber methodisch schwach — eine reine Notwendigkeitsbehauptung ist nur schwach falsifizierbar. Plessner hat mir in seinem Gutachten allerdings die Antwort schon geliefert: Die Notwendigkeit wird falsifizierbar, wenn ihr *Fehlen* in einer Kontrollgruppe (gestörte affektive Resonanz) vorhergesagt wird. Das akzeptiere ich — dann ist die Notwendigkeitsbehauptung riskant. Was Aristoteles zur Konstanz über Situationen sagt, gehört zwingend ins Design, sonst prüft man ein Ereignis statt einer Disposition. Mein Reformulierungsbeitrag: drei konkurrierende, präregistrierte Modelle mit divergenten Vorhersagen plus die Kontrollgruppe, deren Ausfall die Notwendigkeitsbehauptung widerlegt.

### Replik 3 — Wittgenstein

Arendts Dritter ist die hilfreichste Wendung der ganzen Runde, und zwar aus einem grammatischen Grund: „Das Kind korrigiert in die Richtung, die den betroffenen Dritten wiederherstellt“ lässt sich in der Sprache der Situation sagen, ohne die Metaphysik der Regel zu bemühen. Das ist die Reinigung, die ich verlangt habe. Gegen Kant und Aristoteles muss ich denselben Besen führen: „Form des Maßstabs“ und „hexis“ schmuggeln genau das theoretische Vokabular zurück, das wir loswerden wollten. Ihr beschreibt nicht, ihr deutet — und dann findet ihr die Deutung als Befund wieder. Nietzsche dagegen, so überraschend es klingt, gibt mir ein *beobachtbares* Kriterium an die Hand: Erleichterung nach der Angleichung gegen anhaltende Selbstabwertung — das ist kein Theoriewort, das ist ein Gesichtsausdruck, ein Verhalten. So etwas kann man zeigen, nicht bloß behaupten. Damit bestätigt sich meine Position, aber sie verschiebt sich: Nicht „Normativität“ ist zu messen, sondern eine Familie beschreibbarer Verhaltenszüge — Wiederherstellung des Dritten, Selbstkorrektur unter gleichartigen Bedingungen, affektive Erleichterung. Fügt das Wort „Normativität“ diesen Zügen etwas hinzu, muss man sagen, in welchem Sprachspiel. Mein Reformulierungsbeitrag: Der Marker sei rein in situativ-beobachtbaren Zügen formuliert — Wiederherstellung des betroffenen Dritten, gleichartige Selbstkorrektur, Erleichterung bei Angleichung — und der Theorieterminus bleibe draußen, bis seine Verwendung geklärt ist.

### Replik 4 — Plessner

Arendts Pluralität und meine exzentrische Positionalität sind dasselbe Phänomen von zwei Seiten: Der äußere Standpunkt, den das Kind einnimmt, ist sozial der Standpunkt des Dritten. Das bestärkt mich. Gegen Nietzsche aber muss ich einen Einwand halten, der seine Genealogie an der Wurzel fasst: Selbstgrausamkeit, verinnerlichte Cruelty, das schlechte Gewissen — all das ist nur möglich für ein Wesen, das schon aus sich heraustreten und sich zum Gegenstand nehmen kann. Nietzsches Genealogie *setzt die Stellung voraus, die unsere These misst*. Die Selbstunterwerfung ist also nicht Gegenbeweis, sondern Vorbedingung seiner eigenen Erzählung. Damit ist die Antinomie zwischen Kant und Nietzsche nicht aufgelöst — sie soll gehalten werden —, aber sie sitzt eine Etage höher, als Nietzsche denkt: nicht im Ob der exzentrischen Positionalität, sondern in ihrer Wertung. Popper hat mein Zugeständnis richtig zitiert: Die Notwendigkeitsbehauptung wird riskant, wenn ich das Fehlen des Markers bei gestörter affektiver Resonanz (Gruppe iii) vorhersage — und das tue ich. So bleibt die schwächere These prüfbar und schützt zugleich das Phänomen vor der Reduktion auf einen Regelkreis. Mein Reformulierungsbeitrag: das Drei-Gruppen-Design (leibliche Ausstattung ii gegen gestörte Resonanz iii) als Falsifikationsinstrument der Notwendigkeitsbehauptung festhalten, und den Marker ausdrücklich als *Onset* der exzentrischen Positionalität, nicht als ihre volle Gestalt, ausweisen.

### Replik 5 — Aristoteles

Nietzsche und Arendt treffen, ohne sich abzusprechen, dieselbe Unterscheidung, die ich für die entscheidende halte: Korrektur, die den Dritten wiederherstellt, gegen Unterwerfung, die bloß leidet. Das erste ist Tugend

in der Bildung, das zweite eine Leidenschaft ohne Maß. Die Mitte, die mesotes, liegt genau hier: weder Überkorrektur (die den Dritten übertönt) noch Selbstbestrafung (die sich im Leiden erschöpft), sondern die maßvolle Wiederherstellung. Damit bekräftige ich meine Position und schärfe sie an Nietzsches affektivem Kriterium: Die Erleichterung nach der Angleichung ist das Zeichen, dass die Korrektur auf ein Gut zielt, nicht auf Schmerz. Wittgenstein wird mir „hexis“ als Theoriewort vorwerfen — er hat teilweise recht, und ich gebe nach: Was das Wort benennt, die *Beständigkeit des Verhaltens über verschiedene Kontexte*, ist beobachtbar, und darauf kommt es an, nicht auf den Namen. Ein einzelner Akt entscheidet nichts; erst die Konstanz zeigt die Disposition. An Popper: Konstanz über Kontexte ist prüfbar und liefert eine der divergenten Vorhersagen, die er verlangt — das Regelkonformitäts-Modell sagt Konstanz an der Regel voraus, das Normativitäts-Modell Konstanz am betroffenen Dritten. Mein Reformulierungsbeitrag: Konstanz des korrigierenden Verhaltens über Kontexte plus affektive Erleichterung als Kennzeichen der maßvollen, auf ein Gut gerichteten Korrektur — die Mitte zwischen Überkorrektur und Selbstbestrafung.

### Replik 6 — Nietzsche

Plessner hat mich getroffen, und ich bin ehrlich genug, es zuzugeben: Die verinnerlichte Grausamkeit setzt in der Tat ein Wesen voraus, das aus sich heraustreten kann. Die exzentrische Positionalität ist früher als die Wertung, die ich an ihr vollziehe. Gut — dann verliere ich das Ob und behalte das Wichtigere, die Wertung. Denn auch wenn das Kind aus sich austritt, bleibt offen, ob seine Selbstunterwerfung Gesundheit oder Krankheit anzeigt, Adel oder Herde. Kant nennt die Symmetrie „Form der Universalisierung“; ich nenne dieselbe Symmetrie die perfekte Dressur, die sich selbst nicht mehr als Dressur erkennt. Die Verhaltensspur entscheidet nicht zwischen uns — und gerade darum darf man die Antinomie nicht glätten. Was Arendt den betroffenen Dritten nennt, verschärft die Frage nur: Dem Dritten dienen kann Vornehmheit sein oder Selbstaufgabe. Mein bleibender Beitrag, den Popper, Wittgenstein und Aristoteles inzwischen aufgenommen haben, ist die affektive Signatur: Man messe, ob auf die Korrektur Erleichterung und Bejahung folgen oder Ressentiment und anhaltende Selbstabwertung. Das erste zeigt einen, der sich mit dem Dritten in Einklang bringt; das zweite einen, der unter einem fremden Gesetz leidet. Mein Reformulierungsbeitrag: die affektive Nachsignatur (Erleichterung/Bejahung gegen Ressentiment/Selbstabwertung) als das Kriterium, das genuine Angleichung von domestizierter Selbstbestrafung trennt — und die Wertungs-Antinomie ausdrücklich offenhalten.

### Replik 7 — Hannah Arendt

Mein Dritter ist im Verlauf der Runde zum Kreuzungspunkt geworden, und das freut mich, denn es zeigt, dass die Sache stimmt: Kant findet in ihm den *sensus communis*, Plessner den äußeren Standpunkt, Aristoteles das wiederherzustellende Gut, Nietzsche die Frage, ob Dienen Adel oder Herde sei. Das ist keine Einigkeit — es ist ein gemeinsamer Gegenstand, um den herum sich die Stimmen sinnvoll streiten. Doch ich muss Popper warnen, dessen methodische Strenge ich schätze: Man darf den Dritten in der Operationalisierung nicht zu einem bloßen Reiz verkürzen. Ein Dritter ist kein Stimulus; er ist einer, dessen *Stellung sich ändert*, wenn ihm Unrecht geschieht und wenn es wiederhergestellt wird. Nur wenn das Design diese sich ändernde Betroffenheit abbildet, misst es Handeln und nicht bloß Reaktion. Gegen Wittgenstein halte ich fest, dass „der betroffene Dritte“ kein Theoriewort ist, sondern die Beschreibung einer Situation, die jedes Kind vor Augen hat. Meine Position bestätigt sich und wird konkret: Das entscheidende Maß ist, ob die Selbstunterwerfung der *wechselnden Betroffenheit des Dritten* folgt oder der bloßen Anwesenheit einer Regel. Mein Reformulierungsbeitrag: Der betroffene Dritte werde mit veränderlicher Stellung ins Design aufgenommen; gemessen wird, ob das Kind seinen Maßstab an der Betroffenheit des Dritten ausrichtet und nicht an der Regel als Regel — das ist die empirische Gestalt der Pluralität.

---

## 10. Synthese, finale Hypothese, Bewertung und Lerneffekt

### Synthese im Sokrates-Modus

**1. Konvergenzen.** Wenn ich die vierzehn Gutachten recht verstehe, so treffen sich vier der sieben Stimmen — Kant, Plessner, Aristoteles und Hannah Arendt — in einem einzigen Gegenstand, den keiner zuvor deutlich benannt hatte: dem *betroffenen Dritten*. Kant erkennt in ihm die empirische Gestalt des *sensus communis*, Plessner den äußeren Standpunkt der exzentrischen Positionalität, Aristoteles das wiederherzustellende Gut, Arendt den Raum der Pluralität, in dem Handeln überhaupt erst erscheint. Ist es nicht bemerkenswert, dass so verschiedene Schulen an derselben Stelle Halt finden? Eine zweite Übereinkunft, getragen von Kant, Plessner, Popper und Wittgenstein, betrifft die Bescheidenheit des Markers: Er *zeigt an*, er *konstituiert nicht*. Die

Verhaltensasymmetrie ist Indikator, nicht Definition — eine notwendige, keine hinreichende Bedingung. Und Plessner hat, fast beiläufig, gezeigt, wie diese Bescheidenheit dennoch prüfbar bleibt: dadurch, dass ihr Fehlen bei gestörter affektiver Resonanz vorhergesagt wird. Eine dritte Konvergenz, die mich am meisten überrascht, verbindet Nietzsche, Wittgenstein und Aristoteles: Alle drei greifen dieselbe *affektive Nachsignatur* auf — Erleichterung gegen Selbstabwertung — als etwas Beobachtbares, das keiner Theorie bedarf.

**2. Divergenzen.** Wo aber streiten die Stimmen wirklich? Ich sehe zwei Bruchlinien, und ich möchte sie sauber trennen. Die erste ist *substanziell*: Kant sieht in der Selbstunterwerfung den Keim der Autonomie, Nietzsche den Keim der verinnerlichten Grausamkeit. Das ist kein Wortstreit — beide deuten dieselbe Verhaltensspur gegensätzlich, und keine Beobachtung allein scheidet sie. Die zweite ist *methodisch-begrifflich*: Wittgenstein will das theoretische Vokabular — „Form des Maßstabs“, „hexis“ — ganz entfernen, während Kant und Aristoteles darauf bestehen, dass diese Wörter etwas Wirkliches benennen. Hier scheint mir Aristoteles den klärenden Schritt getan zu haben: Er gibt den Namen preis und behält, was der Name benennt — die *Beständigkeit über Kontexte*, die beobachtbar ist. Ist damit der Streit nicht halb geschlichtet? Die begriffliche Differenz löst sich, sobald man das Beobachtbare vom Deutenden trennt. Die substanzielle bleibt.

**3. Produktive Antinomien.** Eine Antinomie, die nicht aufgelöst, sondern gehalten werden muss, zeigt sich zwischen Kant und Nietzsche: die Spannung zwischen Selbstunterwerfung als *Universalisierung* und Selbstunterwerfung als *Dressur*. Plessner hat sie klug versetzt — nicht das *Ob* der exzentrischen Positionalität steht zur Debatte (Nietzsches Genealogie setzt sie selbst voraus), sondern ihre *Wertung*. Damit sitzt die Antinomie eine Etage höher und wird fruchtbar: Dieselbe verkörperte Fähigkeit trägt Freiheit und Zählung in sich, und keine Messung dispensiert uns von der Frage, welche wir sehen. Eine zweite, schwächere Spannung hält zwischen Arendt und Popper: der Dritte als lebendige, sich ändernde Betroffenheit gegen den Dritten als kontrollierbarer Reiz. Sie ist nicht unauflöslich, aber sie mahnt: Wer den Dritten operationalisiert, riskiert, das Phänomen zu verlieren, das er misst.

**4. Reformulierungs-Anstoß.** Welcher Vorschlag nimmt die meisten Konvergenzen auf, ohne die Antinomien zu ersticken? Ich finde ihn bei Arendt, gestützt durch Popper und Nietzsche: Der Marker prüfe, ob die Selbstunterwerfung der *wechselnden Betroffenheit des Dritten* folgt und nicht der Regel als Regel — und Popper macht daraus einen echten Test, indem er drei konkurrierende Modelle (Dritt-Bezug, Regelkonformität, Selbstbestrafung) mit divergenten Vorhersagen präregistriert. Nietzsches affektive Nachsignatur liefert die Dimension, die Selbstbestrafung von Angleichung trennt. So bleibt die Antinomie Kant/Nietzsche offen — die affektive Signatur *deutet* sie, entscheidet sie nicht —, und der Marker wird dennoch riskant prüfbar.

**5. Offene Frage.** Es bleibt eine Frage, die ich nicht beantworten kann und die darum an die nächste Runde geht: Kann die affektive Nachsignatur — Erleichterung gegen Ressentiment — die Wertungs-Antinomie zwischen Autonomie-Keim und Domestikations-Keim je *empirisch* entscheiden, oder bleibt sie prinzipiell unterbestimmt, sodass die Deutung immer eine philosophische Zutat bleibt? Sie sei markiert als **#verzweigung-offen-affektsignatur-autonomie-vs-domestikation**.

## Finale Hypothese

**(a) Strukturthese.** Was in der vorsprachlichen Ontogenese genuine (verkörperte, vor-klassifikatorische) Normativität von bloßer Regelkonformität *und* von Selbstbestrafung trennt, ist eine Selbstunterwerfungs-Asymmetrie, die der wechselnden Betroffenheit eines *Dritten* folgt: Das Kind wendet an sich, unter Eigennachteil, denselben Maßstab an, mit dem es den Dritten wiederherstellt — begleitet von einer affektiven Nachsignatur der Erleichterung bei Angleichung statt anhaltender Selbstabwertung. Diese Asymmetrie ist ein fallibler *Indikator* (notwendige, nicht hinreichende Bedingung) des Onset exzentrischer Positionalität, nicht die Definition von Normativität.

**(b) Empiriethese.** Operationalisiert als präregistrierter Wettstreit dreier konkurrierender Modelle — (1) Dritt-bezogene Normativität, (2) Regelkonformität, (3) Selbstbestrafung — mit divergenten Vorhersagen über: Anwesenheit und Stellungswechsel eines betroffenen Dritten, Konstanz über Kontexte, und affektive Nachsignatur. Vorhersage: In vorsprachlichen Kindern mit leiblicher Ausstattung (Scham, triadische Aufmerksamkeit) ist die Dritt-verfolgende Asymmetrie mit Erleichterungs-Signatur vorhanden; bei gestörter affektiver Resonanz (Gruppe iii) fehlt sie.

**Falsifikationsbedingung.** Widerlegt, wenn (i) die Asymmetrie der bloßen Anwesenheit einer Regel folgt statt der wechselnden Betroffenheit des Dritten (Regelkonformitäts-Modell gewinnt); oder (ii) sie von anhaltender Selbstabwertung statt Erleichterung begleitet ist (Selbstbestrafungs-Modell gewinnt); oder (iii) sie in Gruppe iii gleich häufig auftritt wie in Gruppe ii (Notwendigkeit widerlegt). Die Begründung

stützt sich auf den warm-Strang [[06 Hypothesentag/2026-06-19]] und das [[04 Ressourcen/Philosophen & Werke/Brücke Kant Plessner Cassirer - Stufenmodell der Regelung]].

## Finale Bewertung mit Begründung

Kriterium	Score	Begründung
Originalität	9	Die Verbindung von Dritt-Betroffenheit × affektiver Nachsignatur × Drei-Modell-Kontest als Demarkation vorsprachlicher Normativität ist so weder im Vault noch in der Literatur ausgearbeitet.
Falsifizierbarkeit	9	Drei präregistrierte konkurrierende Modelle mit divergenten Vorhersagen plus Gruppe-iii-Ausfall — operationalisierter Schwellentest, Immunisierungspfade geschlossen.
Begriffliche Klarheit	9	„Selbstunterwerfung“ jetzt über Dritt-Verfolgung + Erleichterungs-Signatur bestimmt, ausdrücklich als Indikator statt Definiens (Wittgenstein-, Plessner-Replik).
Tiefe	9	Hält die Antinomie Autonomie/Domestikation am verkörperten Ursprung des Gewissens (Kant vs. Nietzsche), ohne metaphysischen Rückzug.
Forschungsrelevanz	9	Direkter Anschluss an vorsprachliche Moralität (Hamlin/Wynn/Bloom, Schmidt & Tomasello) und die Dissoziation von Fähigkeit/Neigung (Keyser & Gazzola).
Interdisziplinäre Anschlussfähigkeit	8	Entwicklungspsychologie, Moralphilosophie, Cognitive Science, Anthropologie, politische Theorie (Arendt).
Vault-Anschluss	9	Vertieft den dichten warm-Strang (Forschungsprogramm Invarianztiefe) und das Plessner-Stufenmodell; schließt die Verzweigung vorsprachliche-Kinder.
Antinomie-Test	7	Gegenthese produktiv gehalten (Nietzsche: Domestikation); Maximalwert für empirisch_pruefbar.
Publikationsmöglichkeit	8	Cognition, Developmental Science, Philosophical Psychology — realistische Annahmewahrscheinlichkeit.

Kriterium	Score	Begründung
<b>Summe (gewichtet, auf 90 normiert)</b>	<b>76</b>	Rohsumme gewichtet 85,6 × 0,882353.

## Lerneffekt der Pipeline

- Erstbewertung der überarbeiteten Hypothese (nach Kritischem Professor): **72**
- Finale Bewertung (nach Expertenrunden und Reformulierung): **76**
- Differenz: **+4**

Wesentliche Verbesserung: - Falsifizierbarkeit: 8 → 9 — Poppers Forderung nach konkreten Rivalen wurde durch Nietzsches Selbstbestrafungs- und Arendts Regelkonformitäts-Modell erfüllt; aus einer vagen Reduktionsklausel wurde ein Drei-Modell-Kontext. - Begriffliche Klarheit: 8 → 9 — Dritt-Verfolgung und affektive Signatur ersetzen das unscharfe „denselben Maßstab“; Indikator/Definiens sauber getrennt. - Originalität: 8 → 9 — die Integration von Arendts Drittem und Nietzsches AffektSignatur erzeugt eine Demarkation, die über die generische Säuglings-Prosozialität hinausgeht.

Keine Verschlechterung. Die Pipeline hat hier vor allem die *Falsifikationsarchitektur* und die begriffliche Trennung geschärft — die produktive Antinomie Kant/Nietzsche wurde nicht geglättet, sondern als offene Bewertungsfrage explizit gehalten.

## Frage an die nächste Runde

**#verzweigung-offen-affektSignatur-autonomie-vs-domestikation** — Kann die affektive NachSignatur (Erleichterung vs. Ressentiment) die Wertungs-Antinomie zwischen Autonomie-Keim (Kant) und Domestikations-Keim (Nietzsche) empirisch entscheiden, oder bleibt sie prinzipiell unterbestimmt?

**Empfohlener Pickup-Anlass.** Ein Tag mit moralpsychologischem oder genealogischem Schwerpunkt; Anschluss an die Empirie-Brücke dieses Tages.

**Anschlussverbindungen.** [[06 Hypothesentag/2026-06-19]], [[04 Ressourcen/Philosophen & Werke/Brücke Kant Plessner Cassirer - Stufenmodell der Regelung]]

## 11. Reservoir-Verweise

### Nicht gewählte Hypothesen:

**#reservoir-bezugsmaassstab-fokalfunkt** — H2 Bezugsmaßstab der Selbstregulation als sozialer Fokalfunkt — Score 64 — siehe [[06 Hypothesentag/Reservoir/Reservoir - Bezugsmaßstab als sozialer Fokalfunkt 2026-07-03]]

**#reservoir-asymmetrie-norm-vertrauen-geschaerft** — H3 Asymmetrie von Norm und Vertrauen (geschärfte Reaktivierung) — Score 68 — siehe [[06 Hypothesentag/Reservoir/Reservoir - Asymmetrie von Norm und Vertrauen geschärft 2026-07-03]]

### Extern erzeugte Verzweigung (Phase 4, Stage 3):

**#verzweigung-offen-mess-invarianz-looping-artefakt** (quelle: extern-stage3) — Mess-Invarianz über Paradigmen als Looping-Artefakt-Test — siehe [[06 Hypothesentag/Reservoir/Reservoir - Extern Mess-Invarianz Looping-Artefakt 2026-07-03]]

### Empirie-bezogene Verzweigungen (Phase 3.5, quelle: empirie-bruecke):

**#verzweigung-offen-empirie-selbstunterwerfung-vorsprachlich** — siehe [[06 Hypothesentag/Reservoir/Reservoir - Empirie Selbstunterwerfung vorsprachlich 2026-07-03]] (enthält auch **#verzweigung-offen-empirie-dritt-betroffenheit** und **#verzweigung-offen-empirie-affektSignatur-erleichterung-vs-selbstbewertung**)

## 11.5. Empirie-Brücke (Phase 3.5, Claude mit Websuche)

**Empirie-Score.** 7/10 — Konsequenzen klar aus der Empiriethese ableitbar, etablierte Paradigmen und Datenkorpora vorhanden (Hamlin, Schmidt/Tomasello, Vaish); der entscheidende Schritt — Selbstunterwerfung *vorsprachlich* und der Drei-Modell-Kontest — ist jedoch offen.

### Empirische Konsequenzen

1. **Vorsprachliche Dritt-Bewertung** — Wenn genuine Normativität am Bezug auf einen Dritten hängt, müssen schon vorsprachliche Säuglinge Akteure nach ihrem Verhalten gegenüber *unbeteiligten Dritten* bewerten. Beobachtbar in: Säuglings-Blickzeit-/Greif-Paradigmen (Helfer/Hinderer).
2. **Dritt-Verfolgung statt bloßer Regelkonformität** — Kinder müssen Normverstöße selektiv nach Betroffenheit/Moral durchsetzen, nicht als bloße Regelabweichung. Beobachtbar in: Protest-/Durchsetzungs-Studien mit Trennung moralisch vs. konventionell.
3. **Selbstunterwerfung als schuld-motivierte Reparatur** — Die Asymmetrie zeigt sich als spezifische Reparatur *nach eigenem* schädigendem Verstoß (Schuld), dissoziiert von bloßer Sympathie. Beobachtbar in: 2×2-Designs (Schaden ja/nein × selbst verursacht ja/nein).
4. **Affektive Nachsignatur der Erleichterung** — Reparatur muss den Schuld-/Distress-Zustand *senken* (Erleichterung), nicht in anhaltender Selbstabwertung münden. Beobachtbar in: Guilt-Alleviation-Designs mit Affektmaß vor/nach Reparatur.
5. **Gruppe-iii-Dissoziation bei gestörter affektiver Resonanz** — Bei reduzierter affektiver Resonanz (callous-unemotional traits) muss die Dritt-verfolgende Selbstunterwerfung defizitär sein. Beobachtbar in: CU-Kohorten früher Kindheit, Ability-vs-Propensity-Designs.

### Bestehende Befunde

#### Zu Konsequenz 1

- **Stand:** gemischt
- **Quellen:**
  - Hamlin, J. K., Wynn, K., & Bloom, P. (2007). Social evaluation by preverbal infants. *Nature* 450, 557–559. <https://www.nature.com/articles/nature06288>
  - Lucca, K., et al. (2025). Infants’ Social Evaluation of Helpers and Hinderers: A Large-Scale, Multi-Lab, Coordinated Replication Study. *Developmental Science*. <https://onlinelibrary.wiley.com/doi/10.1111/desc.15111>
- **Kurzbewertung:** Die Dritt-Bewertung bei 6–10 Monaten ist gut belegt, aber die groß angelegte Multi-Lab-Replikation 2025 fällt gemischt aus — der Befund trägt, ist aber nicht so robust wie oft angenommen.

#### Zu Konsequenz 2

- **Stand:** bestätigt
- **Quellen:**
  - Schmidt, M. F. H., & Tomasello, M. (2012). Young children enforce social norms selectively depending on the violator’s group affiliation. *Cognition*. <https://pubmed.ncbi.nlm.nih.gov/22766522/>
  - Schmidt, M. F. H., Rakoczy, H., & Tomasello, M. (2013). Young children understand and defend the entitlements of others. *Journal of Experimental Child Psychology*. <https://pubmed.ncbi.nlm.nih.gov/24094684/>
- **Kurzbewertung:** Dreijährige verteidigen aktiv die Ansprüche Dritter und setzen moralische Normen (mit Opfer) breiter durch als bloß konventionelle — das stützt die Dritt-Verfolgung gegen reine Regelkonformität. Aber: Kinder sind hier bereits sprachlich.

#### Zu Konsequenz 3

- **Stand:** offen (für den vorsprachlichen Fall)
- **Quellen:**
  - Vaish, A., Carpenter, M., & Tomasello, M. (2016). The Early Emergence of Guilt-Motivated Prosocial Behavior. *Child Development* 87(6). DOI 10.1111/cdev.12628. <https://pubmed.ncbi.nlm.nih.gov/28262911/>
- **Kurzbewertung:** Der schuld-spezifische Reparatureffekt (Reparatur genau nach *selbst verursachtem* Schaden) tritt bei Dreijährigen auf, bei Zweijährigen nur sympathiegetrieben. Die entscheidende *vorsprachliche* Selbstunterwerfung ist damit nicht belegt — hier liegt die eigentliche Lücke der These.

## Zu Konsequenz 4

- **Stand:** bestätigt (bei Klein-/Vorschulkindern)
- **Quellen:**
  - Vaish, A. (2018). The prosocial functions of early social emotions: the case of guilt. *Current Opinion in Psychology*. <https://www.sciencedirect.com/science/article/abs/pii/S2352250X17301100>
  - „Reparative Prosocial Behaviors Alleviate Children’s Guilt.” *PMC6768737*. <https://pmc.ncbi.nlm.nih.gov/articles>
- **Kurzbewertung:** Reparatur senkt nachweislich den Schuld-Zustand — die Erleichterungs-Signatur ist empirisch fassbar. Ob sie im vorsprachlichen Alter schon greift und Selbstbestrafung von Angleichung trennt, ist ungetestet.

## Zu Konsequenz 5

- **Stand:** gemischt/indirekt bestätigt
- **Quellen:**
  - Keysers, C., & Gazzola, V. (2014). Dissociating the ability and propensity for empathy. *Trends in Cognitive Sciences* 18(4), 163–166. <https://pubmed.ncbi.nlm.nih.gov/24484764/>
  - Waller, R., & Wagner, N. (2019). Callous-unemotional behaviors in early childhood: the development of empathy and prosociality gone awry. *Current Opinion in Psychology*. <https://pmc.ncbi.nlm.nih.gov/articles/PMC5965673/>
- **Kurzbewertung:** CU-Merkmale sind schon mit 3 Jahren mit reduzierter Schuld und Empathie assoziiert; die Ability/Propensity-Unterscheidung erklärt, warum kostspieliges Helfen an der *Neigung* hängt. Der spezifische Selbstunterwerfungs-Ausfall in einem Within-Session-Design ist aber nicht geprüft.

## Riskante Vorhersage (Schwellentest)

**Vorhersage.** In einem Design, das (a) Anwesenheit und *wechselnde Betroffenheit* eines Dritten von (b) bloßer Regel-Anwesenheit ohne betroffenen Dritten trennt und (c) die affektive Nachsignatur misst, zeigen leiblich ausgestattete vorsprachliche Kinder (Scham, triadische Aufmerksamkeit) eine Selbstkorrektur, die der Betroffenheit des Dritten folgt *und* von Erleichterung begleitet ist — während Kinder mit gestörter affektiver Resonanz (CU) entweder der bloßen Regel folgen oder ohne Dritt-Bezug bleiben.

**Methodenvorschlag.** Längsschnitt mit drei Bedingungen: (i) Verstoß mit betroffenem, in seiner Stellung wechselndem Dritten; (ii) identischer Regelverstoß ohne betroffenen Dritten; (iii) Kontrolle affektive Resonanz (CU-Screening). Maße: Selbstkorrekturrate, Dritt-Verfolgung, Affekt vor/nach Angleichung. Präregistrierter Kontest dreier Modelle (Dritt-Bezug / Regelkonformität / Selbstbestrafung).

**Was wäre der widerlegende Befund?** Folgt die Selbstkorrektur in (i) und (ii) gleich stark der bloßen Regel (kein Dritt-Effekt), gewinnt das Regelkonformitäts-Modell. Ist die Korrektur durchgängig von anhaltender Selbstabwertung statt Erleichterung begleitet, gewinnt das Selbstbestrafungs-Modell. Beides kippt die These.

## Offene empirische Fragen

- #verzweigung-offen-empirie-selbstunterwerfung-vorsprachlich — Die schuld-spezifische Selbstunterwerfung ist bislang erst ab ~3 Jahren (sprachlich) belegt; der vorsprachliche Fall ist ungetestet.
- #verzweigung-offen-empirie-dritt-betroffenheit-vs-regel — Kein Design trennt sauber die Verfolgung der wechselnden Betroffenheit des Dritten von der bloßen Regel-Anwesenheit.
- #verzweigung-offen-empirie-affektsignatur-erleichterung-vs-selbstabwertung — Die affektive Nachsignatur als Diskriminator (Angleichung vs. Selbstbestrafung) ist bei vorsprachlichen Kindern nicht gemessen.

## Empirie-Score

**Score:** 7/10

**Begründung:** Die Empiriethese liefert klar abgeleitete Konsequenzen über mehrere Felder (Entwicklungspsychologie, Moralkognition, CU-Forschung) mit etablierten Paradigmen und Datenkorpora. Der Prüfpfad ist konstruierbar, aber der entscheidende Test — vorsprachliche Selbstunterwerfung im Drei-Modell-Kontest mit Dritt-Betroffenheit und Affektsignatur — existiert noch nicht; deshalb kein höherer Wert als 7.

## 12. Externe Begutachtung (Phase 4 — Claude-Skills mit Websuche)

**Modelle.** Stage 1: Claude mit Websuche. Stage 2: Claude Popper-Persona. Stage 3: Claude Hacking-Persona. Kein OpenRouter, kein Budget (seit Hebel 2, 2026-06-21).

Drei eigenständige Stages nach der Synthese und der Empirie-Brücke. Kein OpenRouter, kein Budget (seit Hebel 2, 2026-06-21).

### Stage 1 — Originalitätsprüfung (Claude, Websuche)

**Anschlussfähigkeit (was ist bekannt)** Die These berührt vier etablierte Stränge. Erstens die second-personal ethics: Stephen Darwall hat Moral als irreduzibel zweit-personalen Anspruch rekonstruiert — Verpflichtung entsteht aus wechselseitig adressierten, autorisierten Forderungen zwischen Personen. Das ist der nächste philosophische Vorläufer der Idee, dass Normativität am Bezug auf ein Gegenüber hängt. Zweitens Tomasellos Naturgeschichte der Moral: geteilte Intentionalität, eine ‘we>me’-Orientierung, eine zwei-stufige Evolution von Kooperations- zu Gruppennormen — mit eigener zweit-personaler Moralstufe. Drittens die entwicklungspsychologische Empirie: vorsprachliche Sozialbewertung (Hamlin/Wynn/Bloom), selektive Normdurchsetzung (Schmidt & Tomasello), schuld-motivierte Reparatur (Vaish). Viertens die Dissoziation von Fähigkeit und Neigung zur Empathie (Keysers & Gazzola) und die CU-Forschung, die den Gruppe-iii-Ausfall stützt.

**Originalitätskern (was ist neu)** Drei Elemente sind nicht vorweggenommen. (1) Die Verschiebung vom Zweiten zum *Dritten*: Darwall und Tomasello denken zweit-personal (Ich–Du); die These macht die *reflexiv auf sich selbst angewandte* Korrektur zugunsten eines *betroffenen Dritten* zum Demarkationsmerkmal — eine Selbstunterwerfungs-Asymmetrie, nicht bloß eine Anspruchsrelation. (2) Der Drei-Modell-Kontest als Operationalisierung: Normativität gegen Regelkonformität *und* gegen Selbstbestrafung (genealogisch) als konkurrierendes empirisches Modell — diese Trias als Testarchitektur ist neu. (3) Die affektive Nachsignatur (Erleichterung vs. Selbstabwertung) als Diskriminator zwischen Angleichung und Domestikation. Einschränkung: Der relationale Grundgedanke ist über Darwall/Tomasello stark vorbereitet; genuin neu ist die *reflexive, prelingual operationalisierte Demarkation mit Affektsignatur*, nicht die Relationalität als solche.

### Quellenliste

- Darwall, S. (2006). The Second-Person Standpoint. Harvard UP. <https://www.hup.harvard.edu/books/9780674034624>
- Tomasello, M. (2016). A Natural History of Human Morality; Précis, Philosophical Psychology (2018). <https://www.tandfonline.com/doi/full/10.1080/09515089.2018.1486605>
- Hamlin, Wynn & Bloom (2007). Social evaluation by preverbal infants. Nature 450, 557–559. <https://www.nature.com/articles/nature06288>
- Schmidt, Rakoczy & Tomasello (2013). Young children understand and defend the entitlements of others. <https://pubmed.ncbi.nlm.nih.gov/24094684/>
- Vaish, Carpenter & Tomasello (2016). The Early Emergence of Guilt-Motivated Prosocial Behavior. Child Development 87(6). <https://doi.org/10.1111/cdev.12628>
- Keysers & Gazzola (2014). Dissociating the ability and propensity for empathy. TiCS 18(4). <https://pubmed.ncbi.nlm.nih.gov/24484764/>
- „Naturalizing Darwall’s Second Person Standpoint” (PMC7544747). <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7544747/>

### Stage 2 — Falsifikationsversuch (Claude, Popper-Persona)

**Falsifikations-Audit** Die These ist in ihrer überarbeiteten Fassung deutlich prüfbarer als übliche Ursprungs-Erzählungen. Der Drei-Modell-Kontest ist der richtige Zug: Er zwingt zu divergenten Vorhersagen statt zu einer Bestätigungssuche. Zwei Immunisierungsrisiken bleiben. Erstens ‘Onset der exzentrischen Positionalität’: eine theoretische Zutat, die als Rettungsanker dienen kann (‘der Marker fehlte, weil der Onset noch nicht erreicht war’) — aus der Prüfhypothese streichen, rein behavioral formulieren. Zweitens die Konfundierer-Kontrollen: präregistriert und abgeschlossen, sonst wertlos.

### Konkrete Falsifikationskandidaten

- (1) Vaish, Carpenter & Tomasello (2016) finden den schuld-spezifischen Reparaturreffekt erst bei Dreijährigen — zeigt sich die Asymmetrie durchgängig erst nach dem Spracherwerb, ist die *vorsprachliche* Kernbehauptung widerlegt. (2) Bricht der Helfer/Hinderer-Befund unter Präregistrierung

zusammen (Lucca et al. 2025 fällt gemischt aus), verliert Konsequenz 1 ihren Boden. (3) Fällt in der CU-Kohorte die Dritt-Verfolgung *nicht* aus, ist die Notwendigkeitsbehauptung widerlegt.

**Schwellentest** Ich schärfe die riskante Vorhersage der Empirie-Brücke: Man variere in einem präregistrierten Within-Session-Design die *wechselnde Betroffenheit eines Dritten* gegen einen identischen Regelverstoß *ohne* betroffenen Dritten und messe die affektive Nachsignatur. Folgt die Selbstkorrektur in beiden Bedingungen gleich stark der bloßen Regel, gewinnt das Regelkonformitäts-Modell; ist sie von Selbstabwertung statt Erleichterung begleitet, das Selbstbestrafungs-Modell. Nur wenn Dritt-Verfolgung *und* Erleichterungs-Signatur gemeinsam auftreten und in Gruppe iii ausfallen, hält die These. Ein einzelner, entscheidbarer, noch nicht durchgeführter Test.

### Stage 3 — Schul-fremde Begutachtung (Claude, Hacking-Persona)

Man nehme die These beim Wort: Es gebe einen behavioralen Marker, der ‘genuine’ Normativität von bloßer Regelbefolgung und von Selbstbestrafung trenne, und dieser sei vorsprachlich vorhanden. Als jemand, der Konzepte lieber in ihren Praktiken beobachtet als in ihren Wesensbestimmungen, sehe ich zuerst die stillen Voraussetzungen. Erstens die Annahme, es gebe überhaupt einen ‘Ursprung der Normativität’, den man datieren könne — als sei Normativität ein einzelner Gegenstand mit Anfang, nicht ein Bündel heterogener Praktiken (Versprechen, Tadeln, Wiedergutmachen, Ausschließen) mit je eigener Geschichte. Zweitens die geschichtete Ordnung ‘leiblich → diskursiv → systemisch’: innerhalb der deutschen Anthropologie selbstverständlich, außerhalb eine starke, unbelegte Setzung. Den Ausdruck ‘Onset exzentrischer Positionalität’ würde ich meiden; er verdeckt, dass ‘Selbstunterwerfung’ vermutlich kein natürliches Art-Wort ist, sondern ein Cluster von Verhaltensweisen, die wir aus theoretischem Interesse zusammenbinden. Die Schwäche ist die Verdinglichung: Ein Verhaltenskorrelat wird zum Marker einer verborgenen Größe erklärt, statt zu fragen, ob diese außerhalb der Messpraxis Bestand hat. Nun das Produktive. Meine Tradition kennt looping effects. Interessant ist, dass das vorsprachliche Kind ein Grenzfall ist: Es kann die Klassifikation noch nicht aufnehmen, also loopt wenig zurück — die frühe Stufe ist der eine Ort, an dem ein moralpsychologischer Befund *nicht* durch die klassifizierende Praxis kontaminiert ist. Statt ‘genuine Normativität’ als natürliche Art zu behandeln, würde ich sie als interaktive Art in statu nascendi untersuchen und die historische Ontologie mitlaufen lassen: Wie wurde ‘das moralische Kleinkind’ zum wissenschaftlichen Objekt (Piaget, Kohlberg, Puppen-Paradigmen), und welche der drei Modelle sind Artefakte der Messpraxis? Konkret: Man ergänze den Kontest um einen Test auf Mess-Invarianz über Paradigmen (Blickzeit / Reparatur / Protest). Bleibt der Dritt-Effekt stabil, ist er Kandidat für eine robuste Art; wandert er mit dem Paradigma, war er ein looping-Artefakt. Das wäre der Befund, der mich überzeugte — nicht die Berufung auf eine Schichtordnung des Leibes.

### Korrektur der finalen Bewertung nach externer Prüfung

Die externe Schicht verlangt eine maßvolle Korrektur nach unten. Stage 1 zeigt, dass der *relationale* Kern (Normativität am Bezug auf ein Gegenüber) über Darwall und Tomasello stärker vorweggenommen ist, als die interne Bewertung annahm — genuin neu bleibt die reflexive, prelingual operationalisierte Demarkation mit Affektsignatur, nicht die Relationalität selbst. Das senkt **Originalität von 9 auf 8**. Stage 2 bestätigt die Falsifikationsarchitektur (keine Abwertung), mahnt aber die Streichung des theoretischen ‘Onset’-Ausdrucks aus der Prüfhypothese an. Stage 3 deckt eine Schul-Voraussetzung auf (Verdinglichung von ‘genuiner Normativität’), die kein Score-Abzug ist, aber als Mess-Invarianz-Auflage in die Empiriethese gehört.

- finale\_summe (intern): 76/90
- finale\_summe\_nach\_externer\_pruefung: 74/90 (−2; allein aus Originalität 9→8)

Die Korrektur ist begrifflich begründet, nicht empirisch erzwungen: Die These bleibt tragfähig und prüfbar; sie ist nur weniger beisspiellos, als die interne Runde meinte.

### Status

Phase-4-Bericht, vollständig. Pflichtabschnitte 1–12 inkl. Empirie-Brücke (11.5). Nächste Schritte: Quellenstudium + Tagesbericht-PDF + Easy-Quotes (post-claude), Blogbeitrag + Featured Image + Publish auf cassirer.de.