Zeitschrift für Anomalistik
Journal of Anomalistics

www.anomalistik.de www.igpp.de

Authors' Response

Response to the Commentaries by Thomas Rabeyron, Harald Walach, and Ian Tierney on the Articles by Maier, Vogel, Storch, and Dechamps (*JAnom*, *26*[1], pp. 14–36; 2026) and Maier, Vogel, and Dechamps (*JAnom*, *26*[1], pp. 37–74; 2026)

Acausal Psychophysical Correlations in a Causally Closed World?

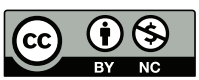
Evidence from (Un)Confounded Erasure Paradigms and the Limits of Physicalist Methodology

Markus A. Maier[a, *], Moritz C. Dechamps[a]

(a) Department of Psychology, Ludwig-Maximilians-Universität, München, Germany

[*] Corresponding author: Markus Maier
markus.maier@psy.lmu.de

PLATINUM OPEN ACCESS
Creative Commons License 4.0
Attribution required.
No commercial use.



Abstract – This response addresses commentaries by Thomas Rabeyron, Ian Tierney, and Harald Walach on two articles in this issue concerning macroscopic complementary relations between subjective and objective aspects of reality, framed within the Generalized Quantum Theory (GQT). The central concern shared by the commentators is whether the erasure-confound paradigm can robustly avoid the non-transmission (NT) axiom-induced decline that has affected earlier experimental paradigms in this research area. We argue that the NT axiom expresses an epistemic restriction arising whenever causal-deterministic documentation is applied to phenomena that contain an autonomous-subjective component. Under this reading, the erasure-confound paradigm achieves only a reduced-objective documentation of the effect at the single-study level, because the confound prevents an unambiguous causal interpretation, while the effect is established indirectly across studies through the exclusion of the alternative explanation (the "color selection

bias" confound). This design decouples evidence accumulation from the arrow of time and is proposed to circumvent the decline. We then put the decline prediction to an empirical test by reporting three previously unpublished color-erasure-with-confound studies (Studies 5–7), which together with the four studies of Maier and Dechamps (2025) form the complete record of seven studies in this paradigm, with no file drawer. The erasure effect appeared in the predicted direction in every study (sign test p = .008), and combined stratified permutation analyses across three independent 36-color sets ruled out the color-selection-bias confound as a joint explanation (Fisher p = .010; Edgington p = .018). An inverse-variance weighted meta-regression on the seven effect sizes returned a non-significant downward slope (b = −0.018, 95% CI [−0.043, 0.006], one-tailed p = .07), and the seventh study yielded an effect still clearly present (BF_{10} = 3,679.73, d = .30). The data are consistent with a stable, reduced-objectively documented acausal psychophysical correlation rather than the attrition predicted by the NT axiom under direct replication. Further specific points raised by individual commentators, such as the experimenter-psi interpretation and the choice of subjective and objective measurements, are addressed in shorter form.

Keywords: Generalized Quantum Theory, non-transmission axiom, NT axiom, macroscopic complementarity, erasure-confound paradigm, psychophysical correlations, decline effect, acausal correlations

We would like to begin by thanking the three commentators (Rabeyron, 2026; Tierney, 2026; Walach, 2026) for their valuable theoretical contributions to our empirical work presented in 2026. We greatly appreciate the time and effort they have invested in evaluating our research program and in initiating a discussion of its theoretical and future empirical implications. We regard supportive and critical arguments alike as the fruitful ground from which a better understanding of meaning-based psychophysical correlations might arise; correlations that we, following the Generalized Quantum Theory (GQT; Atmanspacher, Römer, & Walach, 2002; Römer, 2023; Walach & Römer, 2000), frame as acausal relationships between the subjective realm of reality (volition and experience) and the objective physical aspect of reality (material manifestations).

According to the GQT, the subjective and objective aspects of reality can form macroscopic complementary relations: When both aspects are meaningfully connected by an act of intentional observation, the corresponding psychophysical interaction is proposed to exhibit a number of specific features: (a) it is acausal in nature, meaning that corresponding subjective and objective aspects of reality are not causally related but non-locally correlated or entangled on macroscopic scales; (b) the dual nature of reality is epistemic and therefore observation-dependent, which implies that different forms of observation ranging from passive registration to fully intention-based perception yield different versions of duality, differing in the degree of macro-

scopic entanglement; (c) due to the macroscopically complementary framing of subjective–objective duality, observations of one aspect affect observations of the other, captured by what we have referred to as the incommutability conjecture; and (d) since these acausal psychophysical correlations form subjective–objective quasi-unified realities through entanglement-type correlations, these types of realities cannot be fully objectively documented when subjected to a standard scientific test. This last conjecture is formalized as the non-transmission (NT) axiom, which addresses the paradox involved in scientifically documenting acausal phenomena by means of a causality-testing experimental approach. Its axiomatic status follows from the fact that it cannot be derived mathematically from the GQT formalism. Rather, the NT axiom provides a necessary compensation for the ad hoc omission of Planck's constant to avoid contradictions and paradoxes within a causally closed objective world that scientific documentation needs as a precondition. In our view, however, the NT axiom can be derived exactly from this fundamental dilemma that an objective reality with its causal closure as a precondition for the objectification of phenomena is incompatible with any acausal and thus non-deterministic interference from the subjective realm. Or in other words, meaningful psychophysical correlations constitute sobjective realities and when they are projected epistemically by scientific documentation into the objective world (through causal experimental testing plus direct replication attempts), they gradually (since confirmation through direct replications follows a time signature) lose their sobjective status on the level of scientific confirmation.

The two articles addressed by the three commentators were designed specifically to test the incommutability conjecture and to validate the context derived from the GQT, under which the NT axiom should become active. One article (Maier, Vogel, & Dechamps, 2026) conceptually replicates earlier work (Maier & Dechamps, 2025) using what Rabeyron has aptly termed the "erasure-confound paradigm," here applied to food images with a focus on testing again the incommutability conjecture and the replicability of the erasure results under confound condition. The other article (Maier, Vogel, Storch, & Dechamps, 2026) provides a direct test of the impact of the NT axiom by employing a color-erasure-without-confound paradigm (the proposed context for NT to become active), conceptually replicating but extending the earlier color-erasure-with-confound research of Maier and Dechamps (2025).

Each commentary highlights important aspects of our work and offers ingenious interpretations, alternative theoretical perspectives, and pressing unresolved questions about the meaning of our results, about the theoretical background of the research, and about possible alternative interpretations of the data. Each author provides his individual perspective, but several common concerns also emerge, most prominently, predictions about the long-term reproducibility

of the effects obtained with the erasure-confound paradigm. We do not attempt to address every individual point raised in the commentaries here. Many of the arguments and theoretical interpretations made are highly creative and stand on their own; they reflect the authors' deep familiarity with the GQT, the Pauli–Jung conjecture, and related approaches to reality, and they offer fresh and innovative extensions of these frameworks. We share many of their views and appreciate their constructive engagement with our research program. Nor can we provide answers to all the open questions raised in the comments, partly because we simply do not know the answers and could only speculate, and partly because some questions require additional empirical data before definitive statements about the status of certain GQT-conjectures can be made.

In the present response, we focus on what we take to be the central theoretical concern shared by all three commentators, before turning to a complementary empirical contribution.

The Central Concern

The most central issue raised by Rabeyron (2026) and Walach (2026), and engaged with in a more nuanced way by Tierney (2026), is whether the NT axiom will eventually catch up with the erasure-confound paradigm. The question, in short, is this: Can the erasure-confound design robustly avoid the destructive consequences of the NT axiom, or will studies employing this design, like all previous experimental paradigms in this area, ultimately exhibit the same gradual decline that has plagued the matrix experiments (Walach et al., 2022), the earlier erasure paradigm (Maier et al., 2022), and a long line of related research within experimental parapsychology and anomalistics?

Behind this question lies a deeper one about what the NT axiom actually forbids, and about what "effect documentation" means within the NT axiomatic framework. Did we, in the erasure-confound studies, really confirm a true acausal effect, or did we merely document it indirectly by ruling out the only alternative explanation, leaving the effect itself sufficiently obscured that the NT axiom was not in fact activated? In what follows, we offer theoretical arguments for why the erasure-confound paradigm may indeed be capable of avoiding the activation of the NT axiom at the level of effect documentation. We then provide additional empirical data from three previously unpublished color-erasure-with-confound studies, which together with the four published color studies (Maier & Dechamps, 2025) allow us to put the decline predictions advanced by Rabeyron and Walach to a direct empirical test (Rabeyron, 2026; Walach, 2026).

Reflections about the NT Axiom

To frame our argument, it is useful to begin by reconsidering what the NT axiom is meant to capture. As Tierney notes in his comment (Tierney, 2026), we have suggested in personal communication (and the formulation is worth restating here) that the NT axiom is "not an axiom but naturally follows from the fact that acausal phenomena as proposed by the GQT cannot be robustly documented by a causal, objective approach." On this reading, the NT axiom is not an additional theoretical posit; rather, it expresses the epistemic restriction that arises whenever an attempt is made to test acausal macroscopic relations using a causal testing strategy. The following account of this restriction is necessarily speculative, but it is, we believe, broadly compatible with the framework developed in the Model of Pragmatic Information (MPI; von Lucadou, Römer, & Walach, 2007) and in the GQT (Atmanspacher et al., 2002; Römer, 2023; Walach & Römer, 2000).

In its original MPI formulation, the NT axiom captures a complementary relationship between two components of pragmatic information that arise in a system subjected to scientific documentation: novelty (the autonomy of the phenomenon in the system) and confirmation (the reliability of its documentation). The two components are inversely related: as confirmation/reliability increases through repeated empirical documentation, the novelty/autonomous impulse that originally generated the phenomenon must decrease. This is the formal source of the decline effects observed when acausal phenomena are subjected to direct replication.

More generally, the NT axiom protects the classical, causal-deterministic, and objective description of the world from documentations of phenomena that would contradict it. The classical-causal-deterministic and objective worldview is represented by physicalism, and it is precisely this view that the NT axiom protects with the implication that acausal phenomena of autonomous origin cannot be used for classical signal transfer. If they could, the nomological determinism that underpins, for instance, special relativity would be violated; hence the name non-transmission. The NT axiom thus ultimately protects the causal closure of a physicalist description of the world.

This causal closure is what enables the documentation of physical phenomena in objective form in the first place. The central recognized method of objectification is direct replication: the repeated identical empirical documentation of a phenomenon under largely identical causal conditions. One finds the same causally evoked phenomenon under the same conditions, a state of affairs that is possible only because everything is nomologically determined; otherwise, one would not consistently find the same experimentally produced phenomenon under the same conditions. This, in turn, generates causal closure: everything must be nomologically determined in order to be documentable in this way. The empirical criterion for the demonstration of objective reality therefore presupposes causal closure and determinism.

However, autonomous phenomena, i. e., phenomena that generate genuinely novel and unpredictable states, where "autonomy" is understood in the sense of Kane (2002) as the free origination of a decision without prior causal determination, contradict both nomological determinism and causal closure. In the case of meaningful psychophysical correlations in which an autonomous subjective experience, or an autonomous impulse in the sense of agency, interacts with the objective world, an inherent tension emerges. The autonomy located in (or proceeding from) the subjective realm conflicts with the structure of objective reality and its empirical form of documentation through direct replication. A causal description, or experimental operationalization, of such a psychophysical correlation must therefore itself exhibit a complementary relationship between autonomy and objectification. A high degree of autonomy cannot be objectified through direct replication, because the same thing simply cannot deterministically reappear under the same conditions, which is, after all, the defining feature of autonomy. Conversely, any attempt at objectification through direct replication would fail to capture the autonomy and would therefore declare it non-existent. A direct replication strategy applied to autonomous phenomena will, as the NT axiom predicts, produce a decline of the effect.

This is, in essence, the situation that Rabeyron (2026) describes when he writes that the very conditions under which psi interactions emerge, described as a "communion" or entanglement between the subject and its environment, are invalidated by the standard scientific framework, which presupposes a radical separation of subject and object. The incompatibility between a psychophysical correlation that contains both autonomous-subjective and deterministic-objective components leads during scientific testing to an initial effect documentation followed by a decline, a pattern that is phenomenologically indistinguishable from a false-positive finding that naturally regresses to the mean. This is precisely how the physicalist worldview is preserved: the phenomenon does appear initially, but it can later be declared a false positive and thus ignored. The NT axiom therefore formally guarantees causal closure at the epistemic level, while autonomous psychophysical interactions may continue to exist at the ontic level. In sum, when investigating specific forms of psychophysical correlations in the sense of unified interactions between autonomous-subjective and causally determined-objective aspects of reality, one is confronted with a fundamental obstacle: the incomplete objectifiability of this class of phenomena.

The Loophole: Reduced Objectification Does Not Trigger the NT Axiom

The central point of our argument lies precisely in this notion of incomplete objectifiability, or, in our preferred terminology, reduced-objective documentation. The NT axiom does not, as far as we can see, prohibit reduced or incomplete objective confirmation, so long as the

empirical data do not warrant an unambiguous interpretation in terms of a complete objective demonstration of the phenomenon. In the case of the erasure-confound studies, even a stable replication of the effect at the level of the individual study does not constitute a complete objective demonstration of an erasure effect, because the confound prevents any unambiguous causal interpretation at the single-study level. The subsequent statistical exclusion of the confound through permutation analyses does not alter this situation, because the requirements for an unambiguous, single-study demonstration of a clean erasure effect were never met to begin with. We therefore expect these effects to be robust over time, both within and across studies.

More generally, only methodologically clean individual studies combined with successful direct replication constitute the gold standard of objectification and thus the empirical basis for confirming an effect. Once these criteria are met only in a reduced form, as is the case in the erasure-confound paradigm by design, the complementary destructive effect from the objectification side does not, we propose, fully engage. This is, at least, our hope. Several additional considerations, however, might support it.

Subject–Object Communion, Autonomy, and Two Levels of Measurement

We share Rabeyron's (2026) analysis of psi phenomena and their behavior under scientific testing, and we extend it to the meaning-based psychophysical correlations investigated in the present studies. Rabeyron (2026) correctly notes that psi phenomena emerge from a subject–object communion or entanglement, and that scientific testing requires the separation of this communion, which is precisely the problem. We agree, and we take this to be exactly what the complementarity captured by the NT axiom expresses. It is important, however, to emphasize that the "subject" in this formulation must be understood as an autonomous subject. If subjectivity were construed merely as a mental mirror of the objective world (as it is, for instance, in analytic idealism or epiphenomenalist physicalism) the breaking of the communion would pose no problem at all, because subjectivity and objectivity would simply be two reflections of one and the same phenomenon. Only if subjectivity differs from objectivity in at least one essential respect – and that respect can only be autonomy, since autonomy is the one feature that is not causally determined – does the separation become problematic for scientific documentation. The mind–matter problem is a genuine problem only if autonomy is attributed to the mind. This point is, in our view, captured implicitly by the NT axiom and by the broader GQT and MPI frameworks.

With this in mind, two distinct measurement levels must be distinguished in the psychophysical correlations investigated in Maier and Dechamps (2025) and in the present studies (Maier, Vogel, Storch, & Dechamps, 2026; Maier, Vogel, & Dechamps, 2026). Here, "measure-

ment" is to be understood in the broadest sense as a transfer from unknown to manifest variables, encompassing both perception and the scientific documentation of an effect.

The first measurement level is the individual subjective perception of the color or food stimulus by the participant. When this perceptual act is permeated by individual motives and intentions, e.g. via evaluative judgments concerning how much one likes a color or how appetizing one finds a food image, it constitutes an induced correlation in the Pauli–Jung framework (see Atmanspacher & Rickles, 2022). The perceptual act then contains meaning, sense-making, and an autonomous contribution by the perceiving subject. If, by contrast, the perceiving person were to register the world only passively, without any autonomous willing, the result would be a structural correlation. To ensure the engagement of induced correlations and thereby the activation of the autonomy component required to make a macroscopic complementary relation visible at all, we deliberately focused on motive-supported perceptual processes (color likability and food deliciousness, i.e., preference judgments). Subjective brightness assessments or comparable perceptual judgments would, in our view, yield primarily structural correlations, which would commute normally, produce replicable effects, and exhibit no acausal interactions. The first prerequisite for finding any erasure effect is therefore the activation of an autonomous component in the psychophysical system under study, which is a precondition that has not, in our view, been sufficiently appreciated in previous attempts to investigate macroscopic complementary relations. In sum, the autonomous subjective component is connected with its objective counterpart by performing a motivated perception (level 1) of the outside world which establishes a to-be-studied meaningful psychophysical correlation.

The second measurement level is that of the scientific documentation of the effect: the demonstration of an acausal interaction between autonomous-subjective and causally determined-objective aspects of reality, that is, of the induced correlation itself. This is a collective level at which experimenters, and later observers of the results, are involved. At this level, documentation can proceed in either a fully objectifying way without confound, or in a reduced-objectifying way, as in the "with confound" case, where the erasure effect is veiled by the confound.

The Temporal Projection of Acausal Complementarity

What happens during the scientific documentation of such a psychophysical correlation? We agree here once again with Rabeyron's (2026) analysis: at the moment of documentation, the acausal interaction under investigation is projected into the classical, causal-deterministic, and spatiotemporally structured world. We formulated this point in earlier correspondence, and we are grateful to Tierney (2026) for citing the relevant passage verbatim in his comment.

Time, in particular, appears to us to be the central variable for understanding the consequences of the NT axiom. What occurs in documentation is that a timeless complementarity between classically incompatible opposites (autonomy and determined objectivity), which in timelessness pose no contradiction precisely because they do not causally interact, is projected onto the classical arrow of time. This projection produces causal necessitation not by temporal ordering alone but because scientific documentation operates under the precondition of causal closure. Its central method, direct replication, returns a stable result only if the documented events are nomologically determined, as we argued above. The arrow of time is the axis along which this method accumulates evidence, and the method's built-in assumption of determinism is what converts mere temporal order into the causal frame in which autonomy and objectifiability come into conflict. As a consequence, the level of description changes. What was previously timeless and therefore compatible[1] becomes uniformly temporally directed and incompatible, unless precautionary measures are taken.

In the color-erasure-without-confound study (Maier, Vogel, Storch, & Dechamps, 2026), in which no confound veils the causal structure of the erasure effect, the causal structure upon scientific documentation becomes manifest, and the incompatibility between the present autonomy and the determined objectivity comes into force. The original ontic compatibility due to timelessness is transferred into the four-dimensional spacetime with its clear uniform arrow of time so that both aspects, autonomy and objectifiable fixity, are forced onto the same direction. What the framework predicts robustly at this point is a negative claim: a clean causal test of an autonomous phenomenon cannot terminate in a stable, replicable, confirmed effect, because autonomy and objectification cannot both be satisfied under sustained documentation pressure. Instead, we are observing a pattern in the without-confound case consisting of an

1 The compatibility between autonomy and determined objectivity in a timeless realm can be illustrated through a two-arrows-of-time perspective. Along one time-direction, autonomous decisions appear as undetermined, freely chosen events that only establish a causal chain through autonomously produced links. Viewed backwards, the same chain appears fixed, since its outcomes, which are open from the forward perspective, are accessible as already-realized boundary conditions. The fixity at issue here is retrospective settledness, not nomological necessitation. A freely originated decision in the sense of Kane (2002) is, once made, equally settled in this retrospective sense without thereby having been necessitated by prior states and laws. The two descriptions are therefore not contradictory but perspective-dependent, in analogy to how two-state-vector formulations of quantum mechanics (e.g., Aharonov & Vaidman, 2008) allow both forward and backward boundary conditions to characterize the same quantum system. The forward and backward perspectives are not incompatible in themselves; they come into conflict only when both are collapsed onto a single forward arrow of time, as happens during scientific documentation. This is a speculative framing, but consistent with the broader GQT commitment to acausal ontic relations.

initial accumulation of evidence up to a maximum BF_{10} followed by a decline toward moderate evidence for H0. This pattern is consistent with this negative prediction, and is the one we interpret as autonomy being present early and then squeezed out as objective evidence accumulates. We emphasize, however, that this rise-then-decline signature is phenomenologically identical to a false positive, and we do not treat its specific shape as confirmatory.

It is precisely because a clean causal test yields a trajectory indistinguishable from a false positive that the discriminating empirical test must come from the erasure-confound paradigm, where the predicted pattern, directional consistency across independent replications without decline, is not one that a false-positive account would produce.

From the perspective of scientific documentation, the erasure-without-confound data therefore pose no epistemic threat to the causal closure of the physicalist worldview. Autonomy is, on this account, not objectively detectable under a clean causal test, and any trajectory it produces will be dismissible as a false positive. Under strict causal testing of an erasure effect, the evidence trajectory during scientific documentation is therefore one of effect-to-no-effect, corresponding to the arrow of time and to a failed replication, i. e., a "false positive".

The situation is fundamentally different in studies that employ the erasure-confound paradigm. At the level of the individual study, the reduced internal validity introduced by the confound establishes a reduced-objective evidence for the erasure effect, because the confound provides an alternative explanation, amounting, by standard scientific criteria, to a "no effect" at the single-study level. Through documentation, this psychophysical correlation is again projected onto the classical spacetime with its arrow of time. Studies and their repeated replications produce positive findings, and yet the effect itself cannot be convincingly demonstrated; each study remains in the status of "no effect". Only through the subsequent permutation analysis and similar procedures is the confound retrospectively and temporally subsequently excluded. The erasure effect is not directly demonstrated; rather, the only alternative explanation is excluded. The interpretation that there is a genuine erasure effect, i. e., that the effect is confirmed, is then empirical but only indirect, since only the alternative has been ruled out.

In this design, then, the temporal trajectory of effect production runs from no-effect to effect by exclusion, providing an overall reduced-objective demonstration of the macroscopically complementary and entangled phenomenon. From a physicalist perspective, the classical deterministic worldview and its causal closure remain intact, because each individual study has no probative power, and even the final analysis only excludes an alternative without directly showing the erasure effect. The erasure-confound method therefore achieves two things simultaneously: it decouples the direction of evidence accumulation from the direction of time,

thereby circumventing the false-positive problem; and it documents the effect in question only indirectly, through the exclusion of the alternative explanation. This, in our view, is the reason to expect such effects to remain robust in future replications.

We would add a speculative side note. We believe that spacetime as we know it, with its uniform arrow of time, emerges in its familiar form only through the collective-objectifying method of physicalism. If this is right, then time as it figures in scientific testing, that is, as a causal direction of evidence accumulation, is not a fixed feature of reality but a derivative of how the documentation practice is carried out. The order in which observations are made and aggregated should then itself be a manipulable variable, and decoupling test order from physical time, systematically and orthogonally, could be a powerful way to document autonomy-grounded psychophysical correlations and, more generally, psi effects (see also Schweizer, 2026).

Two further observations are worth making in connection with Rabeyron's (2026) analysis. First, his hypothesis that acausal relationships develop on the ground of causal relationships, with the latter serving as the substrate from which the former emerge, aligns closely with our argument. On Rabeyron's reading, the confound is not merely a veil that protects the acausal correlation from premature objectification; it is the very causal substrate, the "copper of ordinary causality" mixed with "the gold of psi," as he memorably puts it, on which the acausal effect can manifest. Our argument and Rabeyron's (2026) reinforce each other from two angles: from the NT axiom side, the confound prevents the direct causal test that would activate the decline; from the side of the phenomenon itself, the confound provides the causal ground from which the acausal correlation can stably emerge.

Second, Rabeyron (2026) raises the powerful objection that "predicting elusiveness amounts to distorting elusiveness," since elusiveness, if it is genuine, cannot be predicted while remaining elusive. We take this concern seriously. Our response is that the erasure-confound paradigm does not predict the acausal phenomenon directly. What is predicted is a directional difference between two conditions, with the source of that difference left ambiguous at the single-study level. What is then documented post hoc is not the acausal effect itself but the exclusion of the alternative causal explanation; the acausal effect is documented indirectly only by elimination, not by direct prediction-and-confirmation. Whether this distinction is sufficient to satisfy Rabeyron's (2026) elusiveness objection in the long run is, of course, an empirical question to which we now turn.

Empirical Evidence Against a Decline

We acknowledge that all of the foregoing is necessarily speculative, although we think it is internally consistent and grounded in the existing theoretical framework. As a final argument that

the NT axiom does not lead to a decline in the erasure-confound paradigm, we wish to bring empirical data to bear.

Rabeyron (2026) has conducted a decline analysis across the four published food-erasure-with-confound studies and has shown that these four studies exhibit a statistically significant declining trend in effect size (Study 1, $d = .66$; Study 2a, $d = .50$; Study 2b, $d = .29$; Study 2c, $d = .26$; weighted meta-regression $p = .0018$, one-tailed). The question is whether this trend is merely chance variation around a stable effect, or whether the food-erasure-with-confound effect is in fact in the process of declining steadily toward zero, as the NT axiom would predict for a paradigm that ultimately fails to circumvent the axiom. Walach, in his comment (Walach, 2026), makes a related and more specific prediction: at four studies the effect may still be visible, but at seven studies it would likely no longer be detectable.

We do not currently have additional food-erasure-with-confound studies available with which to test this prediction in the food domain. However, in the color-erasure-with-confound research, we have conducted, in addition to the four studies published by Maier and Dechamps (2025), three further studies that have not yet been published and which are described here for the first time. These are the only seven color-erasure-with-confound studies that have ever been carried out under our supervision, i.e. there is no file drawer. They therefore allow us to put both Rabeyron's and Walach's predictions to a direct empirical test in the color domain, which was the original paradigm in which the erasure-confound effect was discovered.

The remainder of this response presents these studies and the corresponding overall analyses across all seven color-erasure-with-confound experiments.

Three Additional Color-Erasure-with-Confound Studies (Studies 5–7)

In addition to the four color-erasure-with-confound studies published by Maier and Dechamps (2025; designated here Studies 1–4 of the color-erasure-with-confound series), three further studies were conducted under our supervision. These are the only three additional studies of this paradigm ever carried out, so that the seven studies presented below, four published and three previously unpublished, constitute the complete record of all color-erasure-with-confound experiments conducted under our supervision. No data have been collected but not reported, and no studies have been selectively excluded. We stress this point because the central claim we wish to advance, namely that the effect remains directionally consistent across all seven studies, would lose much of its force if a file drawer or selective reporting could not be ruled out. The complete data and analysis scripts for Studies 5–7 are available at OSF (https://osf.io/r7sqb).

Three design features distinguish Studies 5–7 from the original four-study series. First, they implemented a within-subjects rather than a between-subjects design in order to increase statistical power (Brysbaert, 2019). Second, they used 36 colors per study (18 randomly assigned to the non-erasure condition and 18 randomly assigned to the erasure condition), with all 36 colors presented to every participant. The increase from 6 colors (Studies 1–2) and 12 colors (Studies 3–4) to 36 colors further reduces the explanatory power of the "biased colors" alternative within each individual study. Third, all three new studies used new color sets drawn independently of those used in Studies 1–4 of Maier and Dechamps (2025), and Studies 6 and 7 in turn drew new color sets independently of each other, so that all three new studies tested fully independent stimulus material.

Beyond these modifications, the methodology of Studies 5–7 followed the food-erasure-with-confound studies reported in 2026 (Maier, Vogel, & Dechamps, 2026) closely, with the substitution of color stimuli (defined by the HSL parameters hue, saturation, and lightness) for the food images. The procedure, ethical approval, online data-collection platform, response scale, sequential Bayesian stopping rules, exclusion criteria (data-integrity item and color-blindness item), and the overall three-step analysis plan are therefore not described in detail again; we refer the reader to the methods section of the food paper (Maier, Vogel, & Dechamps, 2026) for the corresponding information. Only the study-specific details and results are reported below.

Study 5 (Squarra, 2025)

Study 5 was conducted as a Bachelor's thesis under the supervision of the first author (Squarra, 2025). Its aim was to conceptually replicate the overall color-erasure-with-confound effect documented by Maier and Dechamps (2025) using a within-subjects design and an extended stimulus set of 36 colors (18 per condition). The 36 colors were selected via a quantum random number generator (QRNG) prior to data collection. Hue varied randomly between 0° and 359°, lightness between 30% and 70%, and saturation was held constant at 100%. The 18 colors of the non-erasure condition were stored in the result file; the 18 colors of the erasure condition were stored only temporarily in a separate file during data collection (which was required for the program to display them) and were permanently deleted after the conclusion of data collection without prior inspection. Each participant rated all 36 colors on a visual analogue scale from "not at all" (0) to "very much" (100) in response to the prompt "How much do you like this color?". The minimum N to begin Bayesian sequential testing was 100, with an a priori stopping rule of $BF \geq 10$ in favor of H1 or H0, capped at approximately 250 participants for this initial sample.

A total of $N = 246$ participants (after exclusion of $n = 2$, one for self-reported color blindness and one for a failed data-integrity check) provided usable data. A one-sided Bayesian

paired-samples t-test (Cauchy[0, 0.1] prior) yielded BF_{10} = 1.34, indicating anecdotal evidence in support of H1 (d = .09; corresponding frequentist one-tailed p = .085). The mean likability rating was higher in the non-erasure condition (M = 50.82, SD = 12.99) than in the erasure condition (M = 50.08, SD = 12.17). The a priori stopping criterion of $BF \geq 10$ was not reached within the targeted sample.

A technical implementation error in Study 5, which was identified only after data collection had been completed, prevented the planned stratified permutation analysis and multilevel-model analysis from being carried out. Stimulus markers for the colors assigned to the erasure condition were not stored, so that the individual subjective ratings could no longer be aggregated by specific (although unknown) colors within the erasure condition. Without these aggregated color-level scores, the confound-exclusion step of the three-step analysis plan could not be performed at the single-study level. For this reason, Study 5 was not continued and has pilot status only. We nonetheless include it here for completeness, since excluding it would amount to selective reporting. Full methodological details and exploratory analyses (response time, age effects) are reported in Squarra (2025).

Study 6: Preregistered Replication

Study 6 was conducted as a preregistered replication of the within-subjects 36-color design tested in Study 5. The technical error of Study 5 was fixed, so that all 36 colors including those in the erasure condition were associated with stable, internally consistent stimulus identifiers that allowed the planned stratified permutation analysis to be performed at the single-study level. The preregistration is available at OSF (https://osf.io/mkv6x).

In brief, the procedure was a conceptual replication of Studies 1–4 of Maier and Dechamps (2025), with the following modifications: (a) a within-subjects design replaced the between-subjects design, with all 36 colors (18 per condition) presented to each participant; (b) the QRNG-generated colors varied randomly in hue (0°–359°) and lightness (30%–70%), with saturation held constant at 100%; (c) participation was restricted to PC, laptop, or tablet, with smartphone use prohibited (in line with prior observations that erasure effects appear weaker on smartphones); and (d) the Bayesian sequential stopping rule was asymmetric, terminating data collection when either $BF_{10} > 10$ in favor of H1 or $BF_{10} < 1/6$ in favor of H0 was reached, or upon reaching a maximum of N = 5,000 participants. The prior was narrow and centered around 0 (Cauchy[0, 0.1]) and the test was one-sided. The preregistered two-step analysis plan consisted of (1) a Bayesian dependent-samples t-test on subjective likability and (2), conditional on Step 1 yielding strong evidence for H1, a permutation analysis randomly reassigning all 36 colors to conditions across 10,000 permutations, with the criterion $p < .05$. A Bayes Factor

Design Analysis conducted at preregistration estimated an expected stopping point at $n = 683$ in the case of an effect and $n = 1{,}628$ in the case of a null finding. A total of 18 colors assigned to the non-erasure condition and another 18 colors assigned to the erasure condition were drawn randomly within the same parameter ranges and either stored to file or permanently erased without prior inspection (see OSF for color parameter details).

A total of 249 participants took part in the study. After exclusion of $n = 3$ participants (for self-reported color blindness or failed data-integrity check), the final analyzable sample consisted of $N = 246$ participants. The preregistered one-sided Bayesian paired-samples t-test yielded $BF_{10} = 15.19$, indicating strong evidence in support of H1 ($d = .18$; corresponding frequentist one-tailed $p = .002$). The mean likability rating was higher in the non-erasure condition ($M = 57.74$, $SD = 10.93$) than in the erasure condition ($M = 56.54$, $SD = 11.48$).

To address the biased-colors alternative at the level of this single study, a stratified permutation analysis was performed in which color assignments to conditions were randomly permuted 10,000 times, with the mean likability difference between conditions as the test statistic. The observed mean difference was 1.21 (stored – deleted); the permutation null distribution had a mean of 0.003 and a standard deviation of 2.64; the proportion of permutations yielding a difference at least as large as the observed difference was $p = .33$. The biased-colors alternative could thus not be ruled out at the single-study level.

The erasure effect of Study 6 was once again in the predicted direction, conceptually replicating the within-subjects pilot finding of Study 5 and the published cross-study finding of Maier and Dechamps (2025). At the single-study level, however, the stratified permutation analysis did not definitively rule out the confound, i.e., the biased-colors alternative. We therefore conducted Study 7 as a direct replication of Study 6 with an independently drawn new set of 36 colors.

Study 7: Direct Replication of Study 6

Study 7 was a direct replication of Study 6 with the sole modification of an independently drawn new set of 36 colors. The procedure, design, sample acquisition, response measures, sequential Bayesian stopping rule, single-study permutation analysis, and all other analyses were identical to those of Study 6. New sets of 18 colors assigned to the non-erasure condition and of 18 colors assigned to the erasure condition were drawn randomly within the same parameter ranges and either stored to file or permanently erased without prior inspection (see OSF for color parameter details). Study 7 was not preregistered, but the protocol, analysis plan, and stopping rule were identical to those of the preregistered Study 6.

A total of 249 participants took part in the study. After exclusion of n = 4 participants (for self-reported color blindness or failed data-integrity check), the final analyzable sample consisted of N = 245 participants. A one-sided Bayesian paired-samples t-test (Cauchy[0, 0.1] prior) yielded BF_{10} = 3,679.73, indicating extreme evidence in support of H1 (d = .30; corresponding frequentist one-tailed $p < .001$). The mean likability rating was higher in the non-erasure condition (M = 58.06, SD = 11.71) than in the erasure condition (M = 56.03, SD = 11.77). The BF trajectory crossed the BF_{10} = 10 threshold at n = 112 and continued to accumulate strong evidence throughout the remainder of data collection.

To address the biased-colors alternative at the level of this single study, the same stratified permutation analysis was applied. The observed mean difference was 2.03 (stored minus deleted), the permutation null distribution had a mean near zero and a standard deviation of 1.88, and the proportion of permutations yielding a difference at least as large as the observed difference was p = .14. The biased-colors alternative was therefore not ruled out at the single-study level.

Study 7 replicated the erasure effect documented in Studies 5 and 6, with the effect again in the predicted direction and with extreme Bayesian evidence. As in Study 6, the single-study permutation analysis did not exclude the biased-colors alternative on its own. This is the expected situation in this paradigm, where no individual study is designed to carry the full inferential burden, and the exclusion of the confound is established across studies rather than within any one of them, as the following overall analysis shows.

Overall Analyses Across All Seven Color-Erasure-with-Confound Studies

We next consider the evidence against the biased-colors alternative (the confound) across the full seven-study series. The three stratified permutation analyses available to us, which are the combined cross-study permutation of Studies 1 to 4 reported by Maier and Dechamps (2025) and the two single-study permutations of Studies 6 and 7 reported above, share an important methodological property. Each tests the biased-colors alternative on a set of exactly 36 colors. In Maier and Dechamps (2025), the 36 colors are distributed across four between-subjects studies (18 colors in the non-erasure condition and 18 colors in the erasure condition, divided across the four studies as 3+3+6+6 colors per condition). In Studies 6 and 7, the 36 colors of each study (18 per condition) are tested within a single within-subjects design. The three resulting permutation tests therefore arise from methodologically comparable analyses applied to three independently drawn 36-color sets, with the same null hypothesis (random color-to-condition assignment as an alternative explanation for the observed mean difference) and a comparable permutation null distribution structure.

The three corresponding permutation *p*-values were .005 for the Maier and Dechamps (2025) combined analysis, .33 for Study 6, and .14 for Study 7. None of the three excludes the alternative as decisively as a single clean study would, which is consistent with our view that the inferential weight in this paradigm rests at the cross-study level. Because the three 36-color sets were drawn independently via quantum random number generation on three separate occasions, the *p*-values are mutually independent and can be combined with standard methods. The Fisher (1932) combined-probabilities method yielded $\chi^2(6) = 16.75$, $p = .010$. The Edgington (1972) additive combination yielded $p = .018$. Both combinations fall below the conventional significance threshold and converge on the conclusion that across the three independent 36-color permutation analyses, the biased-colors alternative can be ruled out as a joint explanation of the observed pattern.

Two further observations support this conclusion. First, the direction of the erasure effect was hypothesis-consistent in every study, with the mean likability rating higher in the non-erasure than in the erasure condition in each of Studies 1 to 7 (see Table 1). Under the null hypothesis of no effect, the probability of seven directionally consistent results in seven independent studies is $0.5^7 \approx .008$ (one-tailed sign test). This test depends only on the direction of the observed differences and not on any of the within-study assumptions concerning normality, variance, or random color assignment, and it also includes Study 5, which could not include the permutation analyses due to the technical error described above. Second, the seven studies constitute the complete record of all color-erasure-with-confound experiments conducted under our supervision; there is no file drawer.

Taken together, the combined permutation analysis across three independent 36-color sets (Fisher $p = .010$; Edgington $p = .018$), the directional sign test across all seven studies ($p = .008$), and the absence of any file drawer, jointly support the conclusion that the consistent appearance of an effect in the predicted direction across seven studies, with independently drawn color sets, cannot be attributed to chance favorable color assignments. The biased-colors confound can be thereby ruled out at the level of the full series, even though no single study rules it out on its own.

Testing the Decline Hypothesis

The remaining empirical question is whether the seven-study sequence exhibits the systematic decline of effect size that Rabeyron (2026) and Walach (2026) predicted on the basis of the NT axiom. To test this prediction, we applied the same analytical approach Rabeyron used for the four food-erasure studies in his comment, an inverse-variance weighted meta-regression of the effect sizes on study order (1–7), using a fixed-effects model. One caveat applies to this whole

analysis and should be stated at the outset. In the erasure-confound paradigm the effect size of each study conflates the erasure effect with the intrinsic likability difference of that study's particular color draw (confound), so the seven d values are not clean estimates of the erasure effect. We nonetheless run the regression on them, because effect size is the quantity in which Rabeyron (2026) and Walach (2026) framed the decline prediction, and the appropriate way to address their argument is on its own terms.

The seven effect sizes entering the regression are: $d = .21$ (Study 1; Maier & Dechamps, 2025), $d = .50$ (Study 2), $d = .37$ (Study 3), $d = .14$ (Study 4), $d = .09$ (Study 5; Squarra, 2025), $d = .18$ (Study 6), and $d = .30$ (Study 7). The regression yielded a slope of $b = -0.018$ (95% CI [−0.043, 0.006]; $z = -1.46$; $p = .07$, one-tailed), indicating that the null hypothesis of no decline cannot be rejected at conventional significance. A complementary exact permutation test on all 5,040 possible orderings of the seven effect sizes yielded a one-tailed $p = .25$ for the observed slope, again well above conventional significance.

Two contextual comparisons help interpret these statistics. First, the magnitude of the estimated slope ($b = -0.018$ per study) is approximately one-eighth of the slope Rabeyron (2026) reports for the four food-erasure studies ($b = -0.138$ per study). Even taken at face value, the rate of decline in the color domain would be several times slower than in the food domain. Second, and independently of the magnitudes, the shape of the seven-study trajectory is not that of a steady decline. The largest effect was obtained in Study 2 ($d = .50$), followed by a gradual reduction across Studies 3 to 5 (.37, .14, .09), and a partial recovery in Studies 6 and 7 (.18, .30). The sequence of consecutive within-pair changes contains three decreases and three increases, again pointing to an irregular rather than a monotonic pattern.

These analyses jointly fail to support the prediction of a systematic decline across the seven color-erasure-with-confound studies. Walach's (2026) specific prediction that at seven studies the effect would no longer be detectable is not borne out by the data: the seventh and most recent study yielded an effect that was clearly present and directionally consistent. The most parsimonious description of the data is therefore a stable effect with irregular variation across studies, attributable to between-study sampling variability compounded by the study-to-study variability of the confound, rather than to a systematic decline.

Two qualitative observations remain worth making. First, the absence of a statistically reliable decline does not by itself refute the broader theoretical claim that the NT axiom must eventually catch up with the erasure-confound paradigm. It does, however, indicate that the catching-up has not yet occurred within seven studies, a meaningful empirical constraint, since seven is the number Walach proposed as a threshold beyond which detectability would be lost.

Second, the decline question and the existence question remain conceptually distinct. The biased-colors alternative is excluded by the combined permutation analyses and the decline question concerns the long-term behavior of an effect whose existence those analyses leave intact. Table 1 summarizes the key methodological features and the main effects of all seven color-erasure-with-confound studies.

Table 1

Overview of All Seven Color-Erasure-with-Confound Studies on Subjective Likability

Study	Source	Design	Colors per cond.	N	BF_{10}	d
1	Maier & Dechamps (2025)	between	3	945	22.28	.21
2	Maier & Dechamps (2025)	between	3	749	1.33×10^{8}	.50
3	Maier & Dechamps (2025)	between	6	562	603.16	.37
4	Maier & Dechamps (2025)	between	6	1,367	10.19	.14
5	Squarra (2025)	within	18	246	1.34[a]	.09
6	this response	within	18	246	15.19	.18
7	this response	within	18	245	3,679.73	.30

Note. Studies 1–3 used two-sided *BF*s (Cauchy[0, 0.1] prior); Study 4 used a one-sided *BF* (Cauchy[0, 0.2] prior, preregistered). Studies 5–7 used one-sided *BF*s (Cauchy[0, 0.1] prior); Study 6 was preregistered, Studies 5 and 7 were not preregistered.

[a] Pilot status; the a priori stopping criterion of $BF \geq 10$ was not reached. See description of Study 5 in the text.

Conclusion

The NT Axiom and the Erasure-Confound Paradigm: An Empirical Verdict

The empirical evidence presented in the second part of this response converges on a clear picture. Across all seven color-erasure-with-confound studies, the subjective likability of colors was higher when the corresponding objective color parameters were stored than when they were erased; this directional pattern held in every single study (sign test $p = .008$). The combined permutation analyses across three independent 36-color sets ruled out the biased-colors

alternative as a joint explanation (Fisher $p = .01$; Edgington $p = .02$). These two results are the confound-robust core of our case, since the sign test depends only on the direction of the differences and the permutation analyses isolate the erasure signal from the biased-colors draw. The decline question posed by Rabeyron (2026) and Walach (2026) was framed in terms of effect size, and we address it on those terms: an inverse-variance weighted meta-regression on the seven effect sizes returned a slope of $b = -0.018$ with a 95% confidence interval [−0.043, 0.006] that includes zero (one-tailed $p = .07$). The trend is slightly downward but does not reach significance, so the systematic decline predicted by the NT axiom, and specifically by Rabeyron (2026) and Walach (2026), is not statistically supported across the seven-study series. However, we attach limited weight to the magnitude of this slope in either direction, because in the erasure-confound paradigm the raw effect size and Bayes factor of each study conflate the erasure effect with the intrinsic likability difference of the particular color draw (=confound). The quantities that speak to the erasure effect net of the confound are the permutation *p*-values and the consistent direction, and both indicate a stable, present effect rather than a vanishing one. Taken together with the parallel evidence from the four food studies (Maier, Vogel, & Dechamps, 2026), the data support our theoretical argument in the first part of this response: the erasure-confound paradigm, by establishing only a reduced-objective documentation of the effect at the level of each individual study, does not appear to activate the NT axiom-driven decline that has plagued previous direct-replication paradigms in this research area.

One natural objection to this conclusion is that the seven color studies are not strictly identical in design, and that the absence of a decline may therefore be an artifact of repeated conceptual rather than direct replication. On this objection, the introduction of new color sets, the transition from between-subjects to within-subjects designs, the use of different stimulus counts per condition (3, 6, 12, or 18 colors), or even small differences in participant pools effectively "resets" the system across studies, prevents the accumulated documentation pressure that would activate the NT axiom, and thereby produces a misleadingly stable empirical pattern. We consider this a valid objection, and our reply is grounded in the structure of our own data. Studies 5, 6, and 7 share an identical design, a within-subjects 18-vs-18 color comparison with 36 randomly drawn colors per study, conducted on adult online samples of approximately 245 participants each ($N = 246$, 246, and 245, respectively). They are not conceptual replications but direct replications of one another. Under the design-variation objection, this three-study sub-sequence, in which the experimental design is held constant, should provide the cleanest test of the NT axiom-driven decline hypothesis and any genuine decline should be visible here.

What we observe in this sub-sequence is not the decline the NT axiom would predict. The effect appeared in the predicted direction in all three direct replications, and the seventh and

most recent study, far from showing the disappearance Walach (2026) anticipated by this point in the series, produced an effect that was still clearly present (BF_{10} well above the preregistered threshold) and directionally consistent. We also observe that the nominal effect sizes (d = .09, .18, .30) and *BF*s rise across the three direct replications. However, these magnitudes are confounded with the biased-colors draw of each study, as Study 7 makes plain: its extreme *BF* coincides with a single-study permutation p of .14, so the size of that evidence cannot be attributed to the erasure effect alone. What the three identically designed direct replications do establish is the absence of the predicted attrition under identical experimental conditions. Across the sub-sequence the effect did not weaken toward non-detectability but remained present and directionally consistent. The fluctuations across the full seven-study series are then best understood as ordinary between-study sampling variability, compounded by the study-to-study variability of the confound, rather than as NT axiom-driven decline. Walach's (2026) specific prediction that the effect would no longer be detectable by the seventh study is therefore not borne out, with the qualification that detectability in this paradigm is itself established jointly, through the combined permutation analysis and the directional sign test, rather than through any single study's effect size.

Experimenter Psi and the Source of the Effect

A second issue raised, in different forms, by Rabeyron and by Tierney concerns the so-called "experimenter psi" (e-psi) interpretation of our findings. Both commentators ask whether the observed effects might originate not in the participants but in the experimenters, whose expectations and intentions could non-classically influence the system. Tierney, in particular, develops this into one of two distinct interpretive scenarios for our results: either (a) the GQT-based reading we have proposed, in which the effect reflects a macroscopic non-classical correlation between subjective and objective measurements within an entangled participant–stimulus system, or (b) an e-psi reading, in which the effect is generated by the autonomous intentional impulse of the experimenters themselves. Tierney notes that hypothetico-deductive methods may not be able to discriminate between these scenarios, and he therefore suggests inductive approaches as a possible way forward.

We share the substance of these concerns and add three observations. First, we agree with Rabeyron that the very question of "where" psi originates, in participants versus in experimenters, is partially an ill-posed question once one has accepted the GQT framework. The ontic order from which acausal correlations emerge is not organized by the spatiotemporal coordinates that would license a clean source attribution. The experimenters, the participants, the

colors, and the apparatus form what Rabeyron (2026) terms a "communion of wills" within an entangled system. Asking where the autonomous impulse "is" within this system is analogous to asking where the wave is located in a wave–particle complementarity. Second, even if one were to grant the e-psi reading, the resulting phenomenon would still be a macroscopic non-classical correlation; what changes between scenarios (a) and (b) is the participant-versus-experimenter location of the autonomous impulse, not the acausal-versus-causal nature of the underlying coupling. Walach makes the same point in his commentary, and we take it to imply that the e-psi reading does not falsify the GQT interpretation but only relocates the autonomy within an already-entangled system. Third, we agree with Tierney that what our data cannot do (and what arguably no causality-testing experimental design can do) is determine which participant-versus-experimenter localization is the correct one. This is, in our view, exactly what should be expected if the GQT and the NT axiom are correct: a clean local-causal source attribution would itself constitute the kind of causal signal extraction that the NT axiom forbids.

The same logic applies to Rabeyron's (2026) Decision Augmentation Theory (DAT) reading of our results, in which the apparent erasure effect is reinterpreted as an unconscious precognitive selection of "good moments" by the experimenters. Within the GQT framework we have argued for, DAT and the macroscopic-complementarity reading are not mutually exclusive theories; they are two causal stories about an acausally connected system, neither of which can be definitively confirmed or refuted by a causality-testing design. The data are compatible with both, which is precisely what an acausal phenomenon should produce when forced into a causal interpretive frame.

On the Choice of Subjective and Objective Measurements

A brief comment on the choice of measurements. As we argued in the first part of this response, the central methodological precondition for the appearance of a macroscopic complementary relation is that the psychophysical correlation under study must contain an autonomous-subjective component. This is why we focused on motive-laden, preference-based subjective ratings, i. e., color likability in the original studies and food deliciousness in the food paper, rather than on purely descriptive subjective judgments such as brightness assessments. Likability and deliciousness are evaluative acts in which the perceiving subject's motives, preferences, and willing co-determine the response, and they thereby instantiate the induced-correlation type in the Pauli–Jung framework. Brightness assessments, by contrast, are descriptive perceptual judgments that primarily reflect a structural correlation between the objective stimulus property and its subjective registration; they should commute approximately normally and not yield an erasure effect.

The empirical contrast between these two classes of subjective measurements is informative. In Maier and Dechamps (2025), the brightness-variation results did not replicate beyond the initial two studies, whereas the likability effect appeared in the predicted direction across all four published studies and across the three additional studies reported here. The autonomy component thus appears to be the gating variable for the appearance of the effect at the first measurement level (individual perception), and the reduced-objective documentation strategy (the erasure-confound paradigm) is the gating variable for its survival at the second measurement level (scientific documentation across studies). On the objective side, the choice of measurements (the parametric color values and the food images' identifiers) was driven by the corresponding constraint: the objective values had to be stored or deleted from a results file without ever being inspected by experimenters or participants, so that the erasure manipulation could be executed cleanly under the conditions assumed by the GQT for non-locality to operate.

What the Data Show, and What They Cannot Show

We close with a summary. Our data cannot demonstrate autonomy directly, nor can they identify the source from which any autonomous impulse arises. What they can do, in conjunction with the parallel findings from the food domain reported in 2026, is constrain the alternative explanatory frameworks under which the observed pattern across the seven color studies and the four food studies could be accommodated. A purely physicalist account, which postulates the causal closure of a world describable through deterministic natural laws and direct-replication-validated regularities, predicts that macroscopic entanglement-based psychophysical interactions of the kind investigated here should not appear at all, or that any apparent appearance should rapidly dissipate under repeated documentation pressure. Neither prediction is supported by the data. To the extent that physicalism makes these specific predictions, the data declare physicalism incomplete in this empirical domain.

This is, we believe, the maximum that can be achieved with a deductive-causal method applied to a phenomenon whose ontic structure is, on the GQT account, inherently acausal. The deeper question asking what positive theoretical framework should replace physicalism given its incompleteness in this domain cannot, in our view, be answered with the same deductive approach. Tierney's (2026) suggestion that inductive methods may be required to characterize the alternative worldviews more positively is well-taken, and the conjecture he develops at the end of his commentary is one promising direction. The GQT in its current form is our preferred candidate framework. It is, however, only one of several candidate frameworks consistent with the data, and its full validation will require evidence of a kind that cannot, by the very logic

of the NT axiom, be delivered by any single deductive study. What our seven-study sequence does deliver is the more modest but nonetheless substantial result that a properly designed erasure-confound paradigm appears to be capable of documenting an acausal psychophysical correlation in a stable and reproducible way, without triggering the decline that has, until now, accompanied every other approach to this class of phenomena.

References

Aharonov, Y., & Vaidman, L. (2008). The two-state vector formalism: An updated review. In J.G. Muga, R. Sala Mayato, & I.L. Egusquiza (Eds.), *Time in quantum mechanics* (Vol. 1, pp. 399–447). Springer.

Atmanspacher, H., & Rickles, D. (2022). *Dual-aspect monism and the deep structure of meaning*. Routledge.

Atmanspacher, H., Römer, H., & Walach, H. (2002). Weak quantum theory: Complementarity and entanglement in physics and beyond. *Foundations of Physics, 32*(3), 379–406. https://doi.org/10.1023/A:1014809312397

Brysbaert, M. (2019). How many participants do we have to include in properly powered experiments? A tutorial of power analysis with reference tables. *Journal of Cognition, 2*(1), 16. https://doi.org/10.5334/joc.72

Edgington, E.S. (1972). An additive method for combining probability values from independent experiments. *Journal of Psychology, 80*(2), 351–363. https://doi.org/10.1080/00223980.1972.9924813

Fisher, R.A. (1932). *Statistical methods for research workers* (4th ed.). Oliver and Boyd.

Kane, R. (2002). Free will, determinism, and indeterminism. In H. Atmanspacher & R. Bishop (Eds.), *Between chance and choice: Interdisciplinary perspectives on determinism* (pp. 371–406). Imprint Academic.

Lucadou, W.v., Römer, H., & Walach, H. (2007). Synchronistic phenomena as entanglement correlations in generalized quantum theory. *Journal of Consciousness Studies, 14*(4), 50–74.

Maier, M.A., & Dechamps, M.C. (2025). Macroscopic complementary relation between subjective observations and objective measurements of colors. *Journal of Anomalistics, 25*(1), 15–60. https://doi.org/10.23793/zfa.2025.015

Maier, M.A., Dechamps, M.C., & Rabeyron, T. (2022). Quantum measurement as pragmatic information transfer: Observer effects on (s)objective reality formation. *Journal of Anomalous Experience and Cognition, 2*(1). https://doi.org/10.31156/jaex.23535

Maier, M.A., Vogel, A., & Dechamps, M.C. (2026). Macroscopic complementarity between subjective and objective food image assessments: A conceptual replication of the color-erasure effect. *Journal of Anomalistics, 26*(1), 37–74. https://doi.org/10.23793/zfa.2026.037

Maier, M.A., Vogel, A., Storch, A., & Dechamps, M.C. (2026). Non-classical correlation between subjective and objective color observations: Change of effect as a function of its empirical documentation. *Journal of Anomalistics, 26*(1), 14–36. https://doi.org/10.23793/zfa.2026.014

Rabeyron, T. (2026). Beyond the principle of elusiveness? Analysis of the experimental results of the erasure-confound paradigm. *Journal of Anomalistics, 26*(1), 75–96. https://doi.org/10.23793/zfa.2026.075

Römer, H. (2023). *Quanten, Komplementarität und Verschränkung in der Lebenswelt: Verallgemeinerte Quantentheorie*. Lit Verlag.

Schweizer, M.T. (2026). *Expectancy and reality: Perspectives on mind-matter-interactions* [Unpublished doctoral dissertation]. Ludwig-Maximilians-Universität München.

Squarra, N.F. (2025). *Beziehungen objektiver und subjektiver Variablen eines psychophysischen Systems: Eine Farb-Erasurestudie* [Unpublished Bachelor's thesis]. Ludwig-Maximilians-Universität München.

Tierney, I. (2026). Macroscopic complementarity – questions and a conjecture. *Journal of Anomalistics, 26*(1), 97–112. https://doi.org/10.23793/zfa.2026.097

Walach, H. (2026). Chasing the wind: Empirical signatures of generalized entanglement. *Journal of Anomalistics, 26*(1), 113–119. https://doi.org/10.23793/zfa.2026.113

Walach, H., Kirmse, K.A., Sedlmeier, P., Vogt, H., Hinterberger, T., & Lucadou, W.v. (2022). Nailing jelly: The replication problem seems to be unsurmountable. Two failed replications of the matrix experiment. *Journal of Scientific Exploration, 35*(4), 788–828.

Walach, H., & Römer, H. (2000). Complementarity is a useful concept for consciousness studies. A reminder. *Neuroendocrinology Letters, 21*(3), 221–232.

Zusammenfassung

Akausale psychophysische Korrelationen in einer kausal geschlossenen Welt? Evidenz aus (nicht) konfundierten Löschungsparadigmen und die Grenzen physikalistischer Methodologie

Diese Erwiderung beantwortet die Kommentare von Thomas Rabeyron, Ian Tierney und Harald Walach zu zwei Artikeln diesem Heft über makroskopische komplementäre Beziehungen zwischen subjektiven und objektiven Aspekten der Realität im Rahmen der Verallgemeinerten Quantentheorie (VQT). Die zentrale gemeinsame Frage der Kommentatoren lautet, ob das Erasure-Confound-Paradigma den vom Non-Transmission-(NT-)Axiom postulierten Effektrückgang, der frühere experimentelle Paradigmen in diesem Forschungsfeld betroffen hat, dauerhaft vermeiden kann. Wir argumentieren, dass das NT-Axiom eine epistemische

Beschränkung ausdrückt, die immer dann auftritt, wenn eine kausal-deterministische Dokumentation auf Phänomene mit einer autonom-subjektiven Komponente angewendet wird. Unter dieser Lesart erreicht das Erasure-Confound-Paradigma auf der Einzelstudienebene nur eine reduziert-objektive Dokumentation des Effekts, weil der Confound eine eindeutige kausale Interpretation verhindert, während der Effekt studienübergreifend indirekt über den Ausschluss der alternativen Erklärung (des „Color-Selection-Bias"-Confounds) etabliert wird. Dieses Design entkoppelt die Evidenzakkumulation vom Zeitpfeil und soll dadurch den Effektrückgang umgehen. Anschließend prüfen wir die Rückgangsvorhersage empirisch, indem wir drei bislang unveröffentlichte Farb-Erasure-with-Confound-Studien (Studien 5–7) berichten, die zusammen mit den vier Studien von Maier und Dechamps (2025) den vollständigen Datensatz von sieben Studien in diesem Paradigma bilden, ohne Schubladenproblem. Der Erasure-Effekt zeigte sich in jeder Studie in der vorhergesagten Richtung (Vorzeichentest $p = .008$), und kombinierte stratifizierte Permutationsanalysen über drei unabhängige 36-Farben-Sets schlossen den Color-Selection-Bias-Confound als gemeinsame Erklärung aus (Fisher $p = .010$; Edgington $p = .018$). Eine inverse-varianz-gewichtete Meta-Regression über die sieben Effektstärken ergab einen nicht signifikanten negativen Gradienten ($b = -0{,}018$, 95%-KI [−0,043, 0,006], einseitig $p = .07$), und die siebte Studie zeigte einen weiterhin deutlich vorhandenen Effekt ($BF_{10} = 3.679{,}73$, $d = .30$). Die Daten sind mit einer stabilen, reduziert-objektiv dokumentierten akausalen psychophysischen Korrelation vereinbar und nicht mit dem vom NT-Axiom unter direkter Replikation vorhergesagten Effektrückgang. Weitere von einzelnen Kommentatoren aufgeworfene Punkte, etwa zur Experimenter-Psi-Interpretation und zur Wahl der subjektiven und objektiven Messungen, werden in kürzerer Form behandelt.

Schlüsselbegriffe: Verallgemeinerte Quantentheorie, Non-Transmission-Axiom, makroskopische Komplementarität, Erasure-Confound-Paradigma, psychophysische Korrelationen, Effektrückgang, akausale Korrelationen