Journal of Anomalistics
Zeitschrift für Anomalistik

www.anomalistik.de www.igpp.de

Macroscopic Complementarity Between Subjective and Objective Food Image Assessments

A Conceptual Replication of the Color-Erasure with Confound Effect

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

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

(b) M. Maier and A. Vogel contributed equally and share first authorship.

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

Abstract – The Generalized Quantum Theory (GQT) proposes that subjective experiences and objective physical measurements constitute macroscopic complementary subsystems related through acausal, non-local entanglement correlations. The non-commutability conjecture of the GQT predicts that measurement within one subsystem modifies the state of the entire system, including the eigenvalues of the complementary subsystem. Maier and Dechamps (2025) provided preliminary evidence for this conjecture by demonstrating that the storage (non-erasure) versus erasure of objective color parameters (hue and lightness) systematically affected subjective likability ratings across four studies (higher subjective likability mean score in the non-erasure compared to the erasure condition). The present research conceptually replicated these color-erasure findings in a different stimulus domain. In four studies (total *N* = 647), participants rated photographs of food dishes on subjective tastiness while the objective food parameters (calorie values and stimulus IDs) were either stored (non-erasure condition) or permanently deleted (erasure condition). As in the color studies, limited stimulus sets were fixedly assigned to conditions within each study, deliberately introducing a stimulus-bias confound to protect the acausal complementary relation from the destructive impact of the non-transmission (NT) axiom. Following a predefined three-step analysis plan, preregistered Study 1 (Step 1) documented a strong erasure effect on tastiness ratings. Participants evaluated food dishes as more appetizing in the non-erasure compared to the erasure condition (BF_{10} = 1309.22, d = .66), and Studies 2a–2c (Step 2) replicated this effect with three independently drawn stimulus sets (BF_{10}s ranging from 10.40 to 26×10^{15}, ds ranging from .26 to .50). In Step 3, preregistered overall analyses including a stratified permutation test (p = .002), a multilevel model (β = 4.72, p = .011), and sensitivity analyses ruled out the biased stimulus assignment alternative. These findings provide cross-domain evidence for the non-commutability conjecture: the documentation of objective stimulus parameters appears to acausally co-relate with subjective evaluations, consistent with macroscopic complementary relations as proposed by the GQT. The implications for substance dualism, physicalism, and dual-aspect monism are discussed.

Keywords: Generalized Quantum Theory, macroscopic complementarity, macroscopic non-local entanglement correlation, non-commutability, erasure paradigm, psychophysical interaction

Introduction

Macroscopic Complementarity and the Psychophysical Problem

The question of how subjective experience relates to the objective physical world – commonly referred to as the psychophysical problem – has occupied philosophers and scientists for centuries (Ruffing, 2021). Within Descartes' (1641) substance dualism, the subjective and objective realms are posited to interact causally, yet numerous scholars have identified a fundamental explanatory gap that such causal accounts fail to bridge (Chalmers, 1995, 1996; Levine, 1983; Nagel, 1974; Shariff et al., 2008). An innovative approach to this long-standing problem has been put forward by the Generalized Quantum Theory (GQT), developed by Walach and Römer (2000, 2011) and Atmanspacher, Römer, and Walach (2002; see also Fach, 2011; Filk & Römer, 2011; Hinterberger & von Stillfried, 2013; Römer, 2023a, 2023b; Lucadou et al., 2007; Walach & von Stillfried, 2011). Rather than postulating causal mechanisms between mind and matter, the GQT proposes that subjective experiences and objective physical measurements constitute macroscopic complementary subsystems whose interrelation is fundamentally acausal in nature.

The concept of complementarity, originally introduced by Bohr (1928) to account for the wave-particle duality in quantum physics, describes a situation in which two observational perspectives are both necessary for a complete description of a phenomenon yet mutually exclusive in their simultaneous determination. The GQT extends this principle beyond the domain governed by Planck's constant, proposing that analogous complementary relationships can emerge between macroscopic subsystems – specifically between subjective experiential states assessed through a first-person perspective and objective physical states measured through a third-person perspective (Römer, 2023b). When both subsystems are meaningfully connected, for instance through an autonomous and intentional subjective perception of the objective stimulus material, the resulting psychophysical correlation takes on the form of a macroscopic complementary relation. Consequently, the observations derived from each subsystem may be incommensurable: determining the state of one subsystem through measurement changes the state of the entire composite system, including the eigenvalues of the complementary subsystem. This constitutes the non-commutability conjecture of the GQT. If both complementary subsystems are spatiotemporally separated, their measurement outcomes are expected to be related through macroscopic non-local entanglement correlations (Römer, 2023b). Such correlations are acausal in that neither subsystem's measurement outcome can be described as exclusively caused by its complementary counterpart, yet both systematically co-relate.

Maier and Dechamps (2025) recently provided preliminary evidence for the existence of entanglement-based macroscopic complementary relations between subjective assessments and objective measurements of color stimuli, with the data indicating non-commutability between subjective ratings of colors and variations in objective color parameter documentation. The goal of the research presented herein is to re-examine the non-commutability conjecture of the GQT by conceptually replicating these original findings. In this experimental variation, we used different stimulus material involving pictures of food items, from which subjective ratings and variations in the documentation of objective parameters (calorie counts, stimulus IDs) were obtained.

The Problem of Causal Testing and the Non-Transmission Axiom

A central challenge for empirical investigations of macroscopic complementarity lies in the acausal nature of the proposed phenomenon. If such meaning-based psychophysical entanglement correlations could be reliably documented through conventional causal experimental designs, they could in principle be exploited for signal transfer, which would violate the constraints imposed by special relativity (Lucadou et al., 2007; Römer, 2023b). To preclude such paradoxical consequences, the GQT incorporates the non-transmission (NT) axiom, which prohibits the robust causal utilization of macroscopic non-local entanglement correlations (Römer, 2023b). As formalized in the Model of Pragmatic Information (MPI; Lucadou et al., 2007), the NT axiom predicts that when macroscopic complementary relations are subjected to a strict causal experimental test, the resulting data will exhibit an initial effect followed by a systematic decline in evidential support – a pattern that is phenomenologically indistinguishable from a false-positive finding. This effect-and-decline pattern reflects the progressive destruction of the acausal complementary relation through continued causal objectification.

The implication for empirical research is profound: strict, internally valid causal tests of meaning-based acausal psychophysical phenomena are, according to the GQT, ultimately self-defeating. The very act of maximizing the causal explanatory power of an experimental design eliminates the acausal entanglement that generates the phenomenon under investigation. This theoretical constraint necessitates the development of alternative testing strategies that reduce the directness of the causal test while still permitting convincing empirical documentation of the hypothesized complementary relations (see Maier & Dechamps, 2025).

Empirical Evidence from Color-Erasure Studies

Maier and Dechamps (2025) provided the first systematic empirical evidence for entanglement-based macroscopic complementary relations between meaningfully related subjective and objective aspects of color assessment. In a series of four studies, they investigated whether the storage or deletion (erasure) of objective color parameters, hue and lightness in the HSL model, influenced participants' subjective evaluations of the colors presented. In each study, limited subsets of randomly selected colors were fixedly assigned to either an erasure condition, in which the objective color parameters were omitted from the results file, rendering them irretrievable after completion of data collection, or a non-erasure condition, in which these parameters were stored and remained accessible to the experimenters. Participants rated the colors on visual analogue scales for subjective likability (and also brightness) without any knowledge of the erasure manipulation.

Crucially, the use of limited, fixed color subsets per experimental condition was an intentional design feature rather than a methodological weakness. By restricting each condition to a small set of colors, any observed difference in subjective evaluations could alternatively be attributed to a coincidentally favorable assignment of more appealing colors to one condition (the "biased colors" argument). This deliberate confound reduced the internal validity of each individual study, thereby attenuating the directness of the causal test. We decided to use this strategy to protect the acausal complementary relation from the destructive impact of the NT axiom. If a genuine erasure effect exists, it should appear consistently across studies despite the use of different color sets in each study. Only when the confound is disentangled in a final overall analysis can the true causal nature of the manipulation be retrospectively established.

This three-step research agenda, consisting of (1) documenting an effect in each individual study, (2) demonstrating replicability across studies with different stimulus sets, and (3) ruling out the stimulus-bias alternative through overall analyses, yielded compelling results for the proposed macroscopic complementary relation and the resulting non-commutability between measurements of subjective likability of colors and erasure-manipulated objective color parameter documentation. In all four studies, participants rated colors as more likable when the objective color parameters had been stored (non-erasure) compared to when they had been erased. The Bayesian evidence for this effect was strong to extreme in each study (Study 1: BF_{10} = 22.28, d = .21; Study 2: BF_{10} = 133.2 million, d = .50; Study 3: BF_{10} = 603.16, d = .37; Study 4 [preregistered]: BF_{10} = 10.19, d = .14). The overall permutation analysis across all four studies, which randomly re-assigned colors to erasure conditions, demonstrated that the probability of obtaining the observed pattern of results through coincidentally favorable color assignments alone was very low (p = .005 across all studies; p = .05 for confirmatory Studies 3 and 4 only). A supplementary multilevel model analysis confirmed the robustness of the erasure effect after controlling for color-specific variance (β = 7.46, p = .006, d = .28) and showed that study-level differences, including a lightness-range confound present in Studies 1 and 2, contributed negligible variance. These results provided preliminary support for the non-commutability conjecture of the GQT: the objective documentation of color parameters appears to non-classically correlate with subjective evaluations of the corresponding colors.

Strict Causality Test and the Effect-and-Decline Prediction

While Maier and Dechamps (2025) argued on theoretical grounds that a strict, unconfounded causal test of the erasure manipulation should produce an effect-and-decline pattern in line with the NT axiom, empirical evidence for this specific prediction was not provided in their original study series. This gap was addressed by Maier et al. (2026), who performed a conceptual replication of the original color-erasure paradigm using a maximized-causality testing approach. In their study involving 6,448 participants, 20 color stimuli were individually generated for each participant via a quantum random number generator and randomly assigned to erasure and non-erasure conditions in a within-subjects design. This procedure eliminated the biased-colors confound entirely, as each participant received a unique, fully randomized set of colors. The resulting design constituted a strict causal test of the erasure hypothesis.

The results closely followed the prediction derived from the NT axiom. A sequential Bayesian analysis revealed that the erasure effect on subjective likability initially accumulated very strong evidence in favor of the alternative hypothesis, reaching a maximum Bayes Factor of $BF_{10} = 39.77$ at $n = 2{,}946$. This initial finding conceptually replicated the results of the confounded studies (higher mean likability ratings in the non-erasure compared to the erasure condition), now under strict causal testing conditions. However, consistent with the NT axiom, the Bayesian evidence subsequently declined dramatically, ultimately yielding moderate evidence for the null hypothesis ($BF_{01} = 6.16$ at $N = 6{,}448$). An exploratory permutation analysis demonstrated that the probability of obtaining such a pronounced effect-and-decline pattern characterized by reaching the $BF_{10} > 10$ threshold and later declining to at least $BF_{10} = 1$ by chance was only 2.90%, indicating that a mere false-positive interpretation of the data was unlikely. Robustness analyses across different Cauchy priors and a non-parametric reanalysis confirmed the stability of the observed pattern.

Taken together, the absence of a decline in the confounded studies and its clear presence in the unconfounded study are fully consistent with the predictions of the NT axiom and the GQT more broadly. These findings suggest that subjective evaluations and objective measurements of colors constitute macroscopic complementary subsystems: their entanglement correlation and the resulting non-commutability can be stably documented when the causal directness of the experimental test is reduced but exhibit the characteristic effect-and-decline signature when subjected to a strict causality test.

Reclassifying Color-Erasure Effects as Induced Correlations

In the original presentation of their findings, Maier and Dechamps (2025) classified the color-erasure effects as structural correlations in the terminology of the Pauli-Jung framework (see Atmanspacher & Rickles, 2022). Structural correlations were defined as psychophysical interrelations that lack a normative, collectively shared volitional impulse and thus emerge through individualized, context-specific observations from an originally entangled state. Because color preferences were considered non-normative perceptual processes that do not involve stable, goal-directed intentional constructions, the authors argued that the resulting psychophysical system should be classified as a structural correlation, which should be robust and replicable under appropriate testing conditions (Römer, 2023b).

However, upon further theoretical reflection, we now consider this classification to be overly restrictive. The subjective evaluation of how much one likes a color or, as in the present research, how appetizing a food item appears is not a purely passive registering of external stimuli. Rather, it constitutes an autonomous, experience-based act of perception in which individual preferences, past experiences, and motivational states actively contribute to the evaluative judgment. In other words, intrinsic meaning is present in this psychophysical correlation. In the language of the Model of Pragmatic Information (Lucadou et al., 2007), such evaluations contain an autonomy-based novelty component ($N > 0$) reflecting the unique, subjective contribution of the perceiver. This novelty component constitutes an individually autonomous contribution to the measurement process that goes beyond passive registration, rendering the underlying psychophysical correlation meaningful.

We therefore reclassify the erasure-dependent effects on subjective likability – both in the color domain and in the food domain investigated here – as induced correlations. In our revised understanding, the defining feature of an induced correlation is not the presence of a collectively shared normative impulse but the presence of an autonomous, experience-based perceptual contribution ($N > 0$) that results in the meaningful formation of a psychophysical reality with pragmatic information greater than zero. This reclassification does not change the fundamental theoretical predictions regarding the erasure effect or the impact of the NT axiom on causal testing. Rather, it highlights the fact that this meaningful psychophysical reality, through its underlying unification of autonomous subjective and deterministic objective aspects, is not entirely objective in nature and thus cannot be fully objectified with a strict scientific causality-testing approach. This revised view also more accurately reflects and aligns with the classification of induced correlations within the broader theoretical Pauli-Jung framework as described by Atmanspacher and Rickles (2022).

The Present Research

The central question motivating the present research is whether the macroscopic complementary relations and the resulting non-commutability between subjective evaluations and objective measurements documented in the domain of color perception (Maier & Dechamps, 2025) can be conceptually replicated with alternative stimulus material. Specifically, we investigated whether the erasure paradigm generalizes to the domain of food perception, where the objective parameters consist of caloric values and stimulus identifiers associated with images of food dishes and the subjective assessment is a rating of how appetizing or tasty the depicted food appears. In this way, the non-commutability conjecture of the GQT is put to a new test.

The food-erasure paradigm shares the essential structural features of the color-erasure paradigm: objective parameters of the presented stimuli (caloric values and stimulus IDs) are either stored in the results file (non-erasure condition) or permanently deleted and rendered inaccessible (erasure condition) after data collection. Participants evaluate the food images on a subjective tastiness scale without any knowledge of the erasure manipulation. As in the color studies, limited subsets of food images are randomly selected from a larger pool and fixedly assigned to the experimental conditions within each study, thereby introducing the same deliberate confound that reduces the internal validity of each individual study and attenuates the directness of the causal test.

The stimulus material consisted of 100 photographs of attractively arranged food dishes, sourced from the recipe section of a German meal delivery service (HelloFresh). These images were selected to represent a broad variety of cuisines and food types. Each image was associated with an objectively determinable caloric value. In the non-erasure condition, the caloric value together with the stimulus identifier was recorded and stored in a file; in the erasure condition, both pieces of information were not recorded and permanently erased after each study was completed. This manipulation constitutes the independent variable of the present research.

The research program was defined a priori and follows the same data-collection protocol (a first successful Study 1 with three consecutive replications: Studies 2a–2c) and three-step testing agenda established by Maier and Dechamps (2025). In Step 1, the erasure effect on subjective tastiness ratings is documented in a first, preregistered study. In Step 2, the replicability and robustness of the effect are demonstrated across three consecutive independent studies with different, randomly selected stimulus sets. In Step 3, an overall analysis addresses the stimulus-bias confound retrospectively and provides a definitive test of whether the observed effects can be attributed to the erasure manipulation itself rather than to coincidentally favorable stimulus assignments to conditions.

Four studies were conducted, matching the number of studies carried out by Maier and Dechamps (2025) in which the original erasure effect was detected. Study 1 was preregistered and served as the initial test of the food-erasure hypothesis. It employed a within-subjects design with five food images per condition (5 vs. 5), presented across 10 separate sessions with at least two hours between sessions to reduce carryover effects. Each participant rated one image per session, completing 10 sessions in total. Studies 2a, 2b, and 2c (hereafter Studies 2a–2c) served as conceptual replications with a slightly modified design: in each of these studies, six food images per condition (6 vs. 6) were presented within a single session, yielding 12 ratings per participant. Each of the four studies used a different, independently and randomly selected set of food images drawn from the pool of 100 images. Across all studies, the same limited-stimulus-set confound was maintained to protect the acausal nature of the phenomenon, following the rationale established in the color-erasure research.

Hypotheses

Based on the theoretical framework of the GQT proposing non-commutability of measurements derived from meaningfully related macroscopic psychophysical subsystems and based on the empirical precedent established by the color-erasure research (Maier & Dechamps, 2025; Maier et al., 2026), the following hypotheses were formulated.

At the level of individual studies (Steps 1 and 2), it was predicted that participants would rate food images as more appetizing in the non-erasure condition compared to the erasure condition. Specifically, higher mean tastiness scores were expected when the objective caloric values and stimulus identifiers were stored than when they were erased. This prediction was tested in a preregistered manner in Study 1 and in a confirmatory manner, subjected to replicability tests, in Studies 2a–2c. The consistent appearance of this effect across four independent studies with four different sets of randomly selected food images would constitute evidence for Steps 1 (effect documentation) and 2 (replicability) of the research agenda.

At the level of the overall analysis (Step 3), it was predicted that the observed erasure effects could not be attributed to coincidentally favorable stimulus assignments. This hypothesis was operationalized through a preregistered stratified permutation analysis across all four studies. Additionally, a multilevel model analysis was planned in which the erasure manipulation was modeled as a fixed effect and individual stimuli (nested within studies) as random effects, with study included as an additional random factor to account for design differences between Study 1 and Studies 2a–2c. As preregistered, the stratified permutation was expected to yield a significant overall result ($p < .05$), and the fixed effect of the erasure condition in the multilevel model was expected to remain significant after controlling for stimulus-specific and study-level variance. Convergence of both analyses was defined as the strongest evidence against the stimulus-bias explanation. Supplementary sensitivity analyses, including leave-one-stimulus-out analyses and alternative p-value combination methods (Edgington, 1972; Fisher, 1932), were also preregistered to assess the robustness of the overall findings.

Overview of Studies and Three-Step Analysis Plan

Following the three-step analysis plan described above, Step 1 tested the existence of an erasure effect under confound conditions in a first study designed to document an initial erasure effect of objective food image data (erasure vs. non-erasure of calorie values and stimulus IDs) on subjective tastiness rating mean scores. This Study 1 was preregistered on OSF (https://osf.io/zn3sj). Three subsequent replication studies (Studies 2a–2c) were then conducted and are reported below to test the replicability of the initial effect with new stimulus material. These studies constituted Step 2 of the analysis plan. The replication studies implemented an almost identical experimental protocol and the same main analysis protocol within each study. The only differences between Study 1 and Studies 2a–2c were the number of food-picture stimuli within each set (5 vs. 5 in Study 1 and 6 vs. 6 in Studies 2a–2c), the study-specific random stimulus selection prior to each study's data collection (different sets for different studies), and the number of sessions performed (10 separate sessions in Study 1 and one single session in Studies 2a–2c). This altogether resulted in four studies, which matched the number of studies reported by Maier and Dechamps (2025). The decision to conduct four studies was determined a priori before any data collection commenced. After completion of all four studies, a preregistered overall analysis was performed and is reported below to rule out the potential impact of the "biased stimulus assignment" confound on the data. This preregistration (https://osf.io/ghkr7) was uploaded before any permutation or other combined data scores were known. In sum, the studies performed here constitute a conceptual replication of the color-erasure effects reported by Maier and Dechamps (2025), closely matching the original three-step data-collection and analysis plan.

Study 1

Methods

Ethical Guidelines

Prior to the start of the online experiment, participants were presented with general information about the study. Participation was entirely voluntary, and the data protection regulations were outlined. Informed consent was obtained through an active confirmation via button press. All data were analyzed anonymously. The study was designed, conducted, and analyzed in accordance with the APA guidelines and received approval from the Ethics Committee of the Department of Psychology at LMU Munich.

Sample

The data were analyzed using Bayesian sequential procedures, which enable cumulative evidence updating as new participants are tested and their data are successively incorporated into the dataset. Sampling continued until a prespecified Bayes factor (*BF*) threshold provided strong evidence for H0 or H1. Specifically, an a priori evidential criterion of $BF \geq 10$ (strong

Bayesian evidence) was defined, and the *BF* was tracked throughout data collection. Data collection was terminated as soon as the *BF* reached ≥ 10 in favor of H0 or H1, and the resulting final *BF*s were then reported.

The participants were recruited primarily through private contacts and social media by LMU students in the context of experimental courses under the supervision of the first author. During the data collection period, the experimenters were not privy to the details of the experimental manipulation or the hypotheses of the study.

In total, 111 participants took part in the study. Out of these, $n = 56$ did not complete all 10 sessions and were consequently excluded, as specified in the preregistration. As also specified in the preregistration, participants additionally were to be excluded if they indicated that their responses were not reliable by answering a data integrity item with "No" ("Hand on your Heart: Did you really work on this study attentively and conscientiously so that we can use your data?"; $n = 0$). For 19 participants, this question was not displayed because they switched devices between sessions, which prevented the item from being shown in the follow-up survey. Since the 10-session design with forced pauses between sessions required a high level of commitment, we decided to retain these participants in the dataset, nonetheless. Thus, the final dataset comprised $N = 55$ participants. Data collection was considered complete at this sample size because the Bayesian analyses reached the prespecified evidential criterion ($BF \geq 10$).

The sample consisted of 38 female, 16 male, and 1 participant who identified as diverse. The mean age of the participants was 37.71 years ($SD = 17.17$; range: 18–72). Data were collected online via PC or tablet. Participation via smartphone was not supported, and participants completed the study in German (95%) or English (5%).

Materials

Food Stimuli Used in the Experiment. Prior to data collection, 10 distinct food images were randomly selected from the stimulus pool; five were randomly assigned to the erasure condition and five to the non-erasure condition, serving as the experimental stimuli. All 10 experimental stimuli were presented to the participants in a within-subjects design. After pre-study random drawing and assignment to conditions, the selected images and their assignment to experimental conditions were kept unchanged for the entire study. That is, each participant saw the same sets of five stimuli in each condition. The stimulus pool from which the 10 experimental stimuli were drawn consisted of 100 food images selected by the student researchers from recipe listings on the German HelloFresh website, depicting a broad range of food types. Each image was linked to an objective calorie value. These images and their associated calorie information served as the objective stimulus attributes that were either stored or erased as a function of experimental condition. Calorie values ranged from 467 to 1,453 kcal (*M* = 793.29 kcal, *SD* = 184.19 kcal).

Stimulus selection and condition assignment were automated server-side prior to data collection via a Python script. The script employed a hardware quantum random number generator (QRNG; EasyQuantis, ID Quantique) to directly generate 10 unique integers in the range of 0 to 99 without replacement, each corresponding to a food image in the stimulus pool. The first five indices in the generated sequence were assigned to the non-erasure condition and the remaining five to the erasure condition; since the QRNG produced the sequence in random order, this positional split constituted a random assignment to conditions. During data collection, the presentation order of stimuli with their pre-assigned conditions was randomized individually for each participant. Calorie values and stimulus IDs were stored only for the five stimuli assigned to the non-erasure condition. They were automatically stored directly in the experimental file and were accessible to the researchers for subsequent inspection after data collection. Their image IDs and calorie values were: 36, 706 kcal; 63, 825 kcal; 70, 854 kcal; 86, 947 kcal; 29, 688 kcal. Corresponding information for the five stimuli in the erasure condition was only temporarily recorded and later deleted after study completion. Since the image selection procedure was conducted via an automated Python script, which stored the objective stimulus information temporarily in a separate file, no researcher inspected them at any time. This file containing the objective data was deleted following the end of data collection, ensuring that the calorie values and image IDs could never be identified after the experiment. Consequently, the objective stimulus parameters (calorie values and image IDs) were not subjected to inspection during the selection process, nor while data collection was being performed, nor can they ever be identified afterwards. Performing the erasure manipulation after the completion of the experiment was considered necessary (since the program needed the information to run the study) and was considered sufficient to obtain an erasure effect, since time is not linear in the GQT (see Römer, 2023b).

The use of limited subsets of food images (5 vs. 5) randomly assigned to each experimental condition, with this selection and assignment being kept constant for all participants, established an alternative explanation for any potential effects to be found in this study. Specifically, subsets of food images favoring the confirmation of H1 could have been selected into the respective condition by chance. Although this aspect of the design undermines a clear causal interpretation of the erasure manipulation, it was intentionally established to stabilize the acausal nature of the complementary relation tested in this and all subsequent replication studies. As mentioned above, this alternative explanation will be addressed in preregistered final overall analyses at the end of the studies section, provided that Steps 1 and 2 were completed successfully.

Subjective Tastiness Ratings. The food stimuli presented to each participant were rated on a 101-point visual analogue scale (VAS) ranging from "not at all" (0) to "very much" (100). Only the endpoint labels were displayed; numeric values were not visible to participants. For each trial, the food image and the rating scale were presented on the same page. Participants were asked, "How delicious does this food look to you?" (German: "Wie lecker sieht dieses Essen für Sie aus?"). Responses were provided using a horizontal grey bar with the labeled endpoints by positioning a cursor on the bar and confirming the response via mouse click. The cursor initially appeared at the midpoint of the scale to ensure a neutral starting position.

Procedure

Participants were provided with a general overview of the study procedure. They were told that the study consisted of 10 sessions in which images had to be evaluated. The 10 experimental stimuli, including their assigned condition, were presented across 10 different experimental sessions, with one image presentation accompanied by the subjective tastiness rating per session. Each subsequent session was accessible only after a minimum two-hour break between sessions. Data collection across separate sessions, rather than presenting 10 trials within a single session, was implemented because this study was part of a larger data-collection project in which one of the concurrent studies required 10 temporally separated sessions. The other studies conducted alongside the study presented here were not related to it and will be reported elsewhere.

Prior to data collection of Study 1, stimulus selection was performed using the QRNG-based procedure described above. During data collection, the presentation order of stimuli across the 10 sessions was randomized individually for each participant using a deterministic seeded pseudorandom number generator (PRNG), producing a fixed permutation of the 10 stimuli for each participant. This ensured that the presentation order was randomized across participants while remaining reproducible for a given participant across sessions.

It is important to note that in this and all subsequent experiments, participants were unaware of the objective data-storage manipulation (erasure vs. non-erasure). In addition, participants could not identify which condition each food image was assigned to, nor were they informed that two different conditions existed.

After a brief introduction to the task, in each session the food picture was displayed in the center of the screen together with the VAS as described above. Participants then rated the perceived tastiness of the food picture ("How delicious does this food look to you?") by moving the slider to the position on the scale that best reflected their evaluation and confirmed their final rating by pressing the corresponding button.

After completing the ratings across all 10 sessions, participants reported their age and gender. Finally, the data integrity item was presented, asking whether participants had worked attentively and conscientiously so that their data could be used in the analysis. Responses were given by selecting either "Yes" or "No."

Subsequently, the results were saved to a data file that contained only the subjective tastiness data in the erasure condition (with the objective calorie values and image IDs having been permanently deleted), whereas it contained both subjective tastiness data and objective food parameters (image ID and calorie value of the presented food) in the non-erasure condition.

Design and Statistical Analysis

The study used a within-subjects design with one independent variable (IV) with two levels: objective food parameters (picture ID and calorie value of the presented food) were either not available in the final results file (erasure condition) or available (non-erasure condition). The statistical analysis consisted of a one-sided Bayesian paired-samples t-test comparing the erasure and non-erasure conditions. The dependent variable (DV) was the mean score of subjective tastiness ratings for each experimental condition.

For Bayesian inference, we used an uninformed Cauchy prior on the standardized effect size centered at 0 with scale $r = .1$ ($\delta \sim$ Cauchy[0, .1]). Data were accumulated sequentially in chronological order of participation (based on completion time), and Bayes factors were tracked throughout data collection.

All materials, experimental design, hypotheses, and planned statistical analyses reported above were preregistered at OSF (https://osf.io/zn3sj).

Results

To test the directed hypothesis that the mean score of subjective tastiness ratings is higher in the non-erasure than in the erasure condition (non-erasure > erasure), a one-sided Bayesian paired-samples t-test was conducted, with condition (non-erasure vs. erasure) as the within-subject IV and the condition-specific mean score of subjective tastiness ratings as the DV. The Bayesian paired-samples t-test ($N = 55$) yielded a final $BF_{10} = 1309.22$ ($t(54) = 4.88$, $p < .001$), indicating very strong evidence in favor of H1 ($d = .66$). The mean tastiness score was higher in the non-erasure condition ($M = 63.15$, $SD = 14.58$, $n = 55$) than in the erasure condition ($M = 54.99$, $SD = 17.18$, $n = 55$).

In addition to the main analysis, a secondary analysis was preregistered for this study. This involved a permutation analysis with random re-assignments of stimulus-based subjective rating mean scores to experimental conditions to rule out the "biased stimuli" confound. Since this analysis is part of Step 3 of the overall analysis plan, it will be reported in the preregistered overall analysis section below (see sensitivity analyses).

Discussion

The primary objective of Study 1 was to find initial evidence for an erasure-dependent effect of objective food image parameters on subjective tastiness ratings of food images. In the main analysis, we tested whether an experimental manipulation of the documentation of objective food parameters – calorie values and stimulus IDs – affected the subjective evaluation of how appetizing the depicted food appeared. A Bayesian paired-samples t-test with experimental condition (erasure vs. non-erasure of objective food parameters) as IV and mean tastiness rating as DV indicated very strong evidence for H1 (BF_{10} = 1309.22, d = .66). The results demonstrated that participants rated food images as more appetizing when the objective food parameters had been stored (non-erasure condition) compared to when they had been erased.

This finding is consistent with and extends the results reported by Maier and Dechamps (2025), who observed that subjective likability ratings of colors were consistently higher in the non-erasure condition than in the erasure condition across all four of their studies. The present study conceptually replicated this pattern in a different stimulus domain: whereas Maier and Dechamps (2025) manipulated the documentation of objective color parameters (hue and lightness) and measured subjective color likability, the present study manipulated the documentation of objective food parameters (calorie values and stimulus IDs) and measured subjective food tastiness. The convergence of findings across these two distinct stimulus domains suggests that the erasure effect on subjective evaluations may reflect a more general phenomenon rather than being restricted to the specific characteristics of color stimuli.

It is important to note, however, that the results observed in Study 1 can also be attributed to the random selection process employed in choosing the food images used as experimental stimuli. Because only five images were assigned to each condition, it is possible that the images selected for the non-erasure condition happened to depict foods that were inherently rated as more appetizing than those in the erasure condition, resulting in a coincidentally favorable stimulus assignment. This alternative explanation was intentionally built into the design to address the acausal nature of the effect under study. The findings can thus be interpreted in two different ways. They can either be seen as initial evidence supporting the non-commutability conjecture derived from the GQT, suggesting that the documentation of objective food parameters non-classically correlates with subjective evaluations of the corresponding food images. Alternatively, since there were only few food images per condition, the data can be explained by a coincidentally biased stimulus assignment to conditions. This alternative explanation is currently equally viable and must be addressed by subsequent independent replication attempts and a final overall analysis. Only if the effect found in Study 1 (Step 1) is consistent across different random stimulus selections in independent studies (Step 2) and overall analyses can exclude the biased stimulus argument (Step 3) can the true nature of the effect be determined. Consequently, Studies 2a–2c were conducted to replicate Study 1 with newly selected sets of food images.

Study 2a

The aim of Study 2a was to replicate the effect reported in Study 1, following the logic of the color-study paradigm described by Maier and Dechamps (2025). The replication applied an almost identical experimental protocol with the following differences. First, Study 2a used 12 food images (instead of 10) to match the number of stimuli employed in Studies 3 and 4 of Maier and Dechamps (2025). The 10-stimulus set in Study 1 resulted from the 10-session design, which was central to another study that was also part of the research project, limiting the number of stimuli in Study 1 to 10. This limitation was absent in Study 2a. Second, a newly and randomly selected set of food stimuli, with two new subsets of six images for each condition, was used. Third, instead of 10 separate sessions with one trial each, all 12 trials were presented within a single session. The main objective was to replicate the experimental effect observed in Study 1, using the mean score of subjective tastiness ratings as the dependent variable. Thus, Study 2a (and subsequent Studies 2b and 2c) constituted Step 2 of the three-step analysis plan presented by Maier and Dechamps (2025), in which the confound-loaded erasure effect on subjective tastiness ratings of food images was tested for replicability across three independent studies.

Methods

Ethical Guidelines

Prior to the start of the online experiment, participants were presented with general information about the study. Participation was entirely voluntary, and the data protection regulations were outlined. Informed consent was obtained through an active confirmation via button press.

All data were analyzed anonymously. The study was designed, conducted, and analyzed in accordance with the APA guidelines and received approval from the Ethics Committee of the Department of Psychology at LMU Munich.

Sample

The same Bayesian sequential analysis procedures as those employed in Study 1 were used to analyze the data, with a minimum evidential criterion of $BF \geq 10$ (strong evidence) serving as the stopping rule in the main analysis. If this criterion was met, data collection was terminated and the final *BF*s were subsequently reported.

Participants were recruited primarily through private contacts and social media by LMU students as part of bachelor's thesis projects under the supervision of the first author. During data collection, the experimenters were not privy to the details of the experimental manipulation or the study hypotheses.

In total, 369 participants took part in the study (completers only). As specified a priori, participants were to be excluded if they indicated that their responses were not reliable by answering the "Hand on your Heart" item with "No" ($n = 6$). Thus, the final dataset comprised $N = 363$ participants. Data collection was considered complete at this sample size because the Bayesian analyses reached the prespecified evidential criterion ($BF \geq 10$).

The sample consisted of 211 female, 139 male, and 13 participants who identified as diverse. The mean age of the participants was 30.68 years ($SD = 9.56$; range: 18–100). Data were collected online via PC or tablet. Participation via smartphone was not supported, and participants completed the study in German.

Materials

Food Stimuli Used in the Experiment. In contrast to Study 1, Study 2a used a total of 12 food images derived from the same stimulus pool as those in Study 1. Prior to data collection, 12 pictures were randomly selected and six of those assigned to the erasure condition and six to the non-erasure condition (93, 1,078 kcal; 18, 633 kcal; 9, 591 kcal; 54, 773 kcal; 49, 756 kcal; 14, 616 kcal). Unlike Study 1, in which the QRNG directly generated the stimulus indices, stimulus selection in Study 2a was performed via a QRNG-generated seed. Specifically, a 32-byte random seed was generated by the Quantis USB hardware QRNG and served to the experimental program via a local server endpoint. The client-side experimental script then converted this seed into 12 stimulus indices of the stimulus pool range (0–99). The first six resulting indices were assigned to the non-erasure condition and the remaining six to the erasure condition. The seed was generated once at server startup prior to data collection and served identically to all participants, ensuring that the same 12 stimuli and condition assignments were used throughout the study. After data collection, the server was shut down, destroying the seed.

As in Study 1, calorie values and stimulus IDs were stored only for the six stimuli assigned to the non-erasure condition. Corresponding information for the six stimuli in the erasure condition was not retained in the output data: on each erasure-condition trial, the actual image path was overwritten with a generic identifier (e.g., “food7”) before being committed to the data file. The seed required to reconstruct the original stimulus-to-condition mapping existed only in volatile server memory (RAM) and was never written to persistent storage. Upon server shutdown after data collection, the seed was irrecoverably destroyed. This represents a methodological improvement over the file-based erasure procedure used in Study 1, where the temporary file containing objective stimulus parameters was deleted from disk after data collection. Although that file was never inspected, deleted files can in principle be recovered from persistent storage using forensic tools. The ephemeral-computing approach employed here eliminates this residual recoverability, ensuring that the objective parameters of erasure-condition stimuli are irretrievable by any means after the server process has been terminated.

Subjective Tastiness Ratings. Each food picture presented to participants was rated for subjective tastiness. Ratings were provided on visual analogue scales ranging from 0 to 100, as in Study 1 (see the corresponding section of Study 1 for details).

Procedure, Design, and Statistical Analyses

The procedure, design, and statistical analyses, including the Cauchy prior, were identical to those employed and preregistered in Study 1, with the following exceptions: (a) all 12 trials were presented within a single experimental session (rather than across 10 separate sessions), (b) the presentation order of conditions across trials was randomized for each participant using a shuffle function, which relies on the browser’s pseudorandom number generator, and (c) stimulus selection was performed with replacement, meaning that the same image could in principle be drawn more than once across the 12 selections (whereas Study 1 enforced sampling without replacement). The minimum n to initiate Bayesian testing in this and all subsequent studies was set to $n = 100$.

Results

To test the directed hypothesis that the mean score of subjective tastiness ratings is higher in the non-erasure than in the erasure condition (non-erasure > erasure), a one-sided Bayesian paired-samples t-test was conducted, with condition (non-erasure vs. erasure) as the within-subject IV and the condition-specific mean of subjective tastiness ratings as the DV. The Bayesian paired-samples t-test ($N = 363$) yielded a final $BF_{10} = 26 \times 10^{15}$ ($t(362) = -9.61$, $p < .001$), indicating very strong evidence in favor of H1 ($d = .50$). The mean tastiness score was higher in the non-erasure condition ($M = 72.12$, $SD = 16.04$, $n = 363$) than in the erasure condition ($M = 67.03$, $SD = 17.62$, $n = 363$).

Study 2b

The aim of Study 2b was to replicate the effects observed in Studies 1 and 2a. The experimental design followed the procedures of the preceding Study 2a exactly. The only difference was that in Study 2b a newly drawn set of 12 food images was used as stimuli. The random selection of all food stimuli prior to data collection followed the same protocol as in Study 2a, and the identical selection procedure was applied in both experimental conditions. The objective of Study 2b was to replicate the confound-loaded experimental erasure manipulation effect on mean subjective tastiness ratings as the dependent variable.

Methods

Ethical Guidelines

Prior to the start of the online experiment, participants were presented with general information about the study. Participation was entirely voluntary, and the data protection regulations were outlined. Informed consent was obtained through an active confirmation via button press. All data were analyzed anonymously. The study was designed, conducted, and analyzed in accordance with the APA guidelines and received approval from the Ethics Committee of the Department of Psychology at LMU Munich.

Sample

The same Bayesian sequential analysis procedures as those employed in Studies 1 and 2a were used to analyze the data, with a minimum evidential criterion of $BF \geq 10$ (strong evidence) serving as the stopping rule in the main analysis. If this criterion was met, data collection was terminated and the final BFs were subsequently reported.

Participants were recruited primarily through private contacts and social media by LMU students in the context of experimental courses under the supervision of the first author. During data collection, the experimenters were not privy to the details of the experimental manipulation or the study hypotheses.

In total, 121 participants took part in the study (completers only). As specified a priori, participants were to be excluded if they indicated that their responses were not reliable by answering the "Hand on your Heart" item with "No" ($n = 0$; data not available for $n = 19$). Thus, the final dataset comprised $N = 121$ participants. Data collection was considered complete at this sample size because the Bayesian analyses reached the prespecified evidential criterion ($BF \geq 10$).

The sample consisted of 52 female, 68 male, and 1 participant who identified as diverse. The mean age of the participants was 28.69 years ($SD = 13.75$; range: 18–77). Data were collected online via PC or tablet. Participation via smartphone was not supported, and participants completed the study in German.

Materials

Food Stimuli Used in the Experiment. In line with Study 2a, Study 2b used a total of 12 independently drawn food images. Prior to data collection, six pictures were randomly selected and assigned to the erasure condition and six to the non-erasure condition (9, 591 kcal; 4, 564 kcal; 28, 683 kcal; 76, 870 kcal; 64, 830 kcal; 82, 926 kcal), serving as the experimental stimuli. The random selection procedure followed the same protocol as in Study 2a.

Subjective Tastiness Ratings. Each food picture presented to participants was rated for subjective tastiness. Ratings were provided on visual analogue scales ranging from 0 to 100, as in Study 1 (see the corresponding section of Study 1 for details).

Procedure, Design, and Statistical Analyses

The procedure, design, and statistical analyses, including the Cauchy prior, were identical to those employed in Study 2a.

Results

To test the directed hypothesis that the mean score of subjective tastiness ratings is higher in the non-erasure than in the erasure condition (non-erasure > erasure), a one-sided Bayesian paired-samples t-test was conducted, with condition (non-erasure vs. erasure) as the within-subject IV and the condition-specific mean score of subjective tastiness ratings as the DV. The Bayesian paired-samples t-test ($N = 121$) yielded a final $BF_{10} = 27.23$ ($t(120) = -3.18$, $p = .001$), indicating strong evidence in favor of H1 ($d = .29$). The mean tastiness score was higher in the non-erasure condition (*M* = 66.17, *SD* = 14.24, *n* = 121) than in the erasure condition (*M* = 63.38, *SD* = 15.31, *n* = 121).

Study 2c

Study 2c was conducted as a final replication test of the confound-loaded experimental erasure manipulation effects on subjective tastiness ratings reported in Studies 1, 2a, and 2b. This study completed the pre-planned series of four studies and constituted the final replicability test of Step 2 within the three-step analysis plan introduced by Maier and Dechamps (2025). The procedure and statistical analyses followed the same protocol as in the previous two studies. Prior to data collection, a newly drawn set of 12 food pictures (six per condition) was randomly selected and used as the experimental stimuli. The dependent variable was the mean tastiness rating score.

Methods

Ethical Guidelines

Prior to the start of the online experiment, participants were presented with general information about the study. Participation was entirely voluntary, and the data protection regulations were outlined. Informed consent was obtained through an active confirmation via button press. All data were analyzed anonymously. The study was designed, conducted, and analyzed in accordance with the APA guidelines and received approval from the Ethics Committee of the Department of Psychology at LMU Munich.

Sample

The same Bayesian sequential analysis procedures as those employed in the preceding studies were used to analyze the data, with a minimum evidential criterion of BF ≥ 10 (strong evidence) serving as the stopping rule in the main analysis. If this criterion was met, data collection was terminated and the final BFs were subsequently reported.

Participants were recruited primarily through private contacts and social media by LMU students as part of bachelor's thesis projects under the supervision of the first author. During data collection, the experimenters were not privy to the details of the experimental manipulation or the study hypotheses.

In total, 111 participants took part in the study (completers only). As specified a priori, participants were to be excluded if they indicated that their responses were not reliable by answering the “Hand on your Heart” item with “No” ($n = 3$). Thus, the final dataset comprised $N = 108$ participants. Data collection was considered complete at this sample size because the Bayesian analyses reached the prespecified evidential criterion ($BF \geq 10$).

The sample consisted of 63 female and 45 male participants. The mean age of the participants was 31.09 years ($SD = 16.34$; range: 18–80). Data were collected online via PC or tablet. Participation via smartphone was not supported, and participants completed the study in German.

Materials

Food Stimuli Used in the Experiment. In line with Studies 2a and 2b, Study 2c used a total of 12 new randomly selected food images. Prior to data collection, six pictures were randomly selected and assigned to the erasure condition and six to the non-erasure condition (90, 969 kcal; 99, 1,409 kcal; 66, 833 kcal; 80, 895 kcal; 52, 766 kcal; 38, 716 kcal), serving as the experimental stimuli. The random selection procedure followed the same protocol as in the preceding studies.

Subjective Tastiness Ratings. Each food picture presented to participants was rated for subjective tastiness. Ratings were provided on visual analogue scales ranging from 0 to 100, as in Study 1 (see the corresponding section of Study 1 for details).

Procedure, Design, and Statistical Analyses

The procedure, design, and statistical analyses, including the Cauchy prior, were identical to those employed in Studies 2a and 2b.

Results

To test the directed hypothesis that the mean score of subjective tastiness ratings is higher in the non-erasure than in the erasure condition (non-erasure > erasure), a one-sided Bayesian paired-samples t-test was conducted, with condition (non-erasure vs. erasure) as the within-subject IV and the condition-specific mean score of subjective tastiness ratings as the DV. The Bayesian paired-samples t-test ($N = 108$) yielded a final $BF_{10} = 10.40$ ($t(107) = -2.75$, $p = .004$), indicating strong evidence in favor of H1 ($d = .26$). The mean tastiness score was higher in the non-erasure condition ($M = 59.99$, $SD = 13.79$, $n = 108$) than in the erasure condition ($M = 56.24$, $SD = 15.25$, $n = 108$).

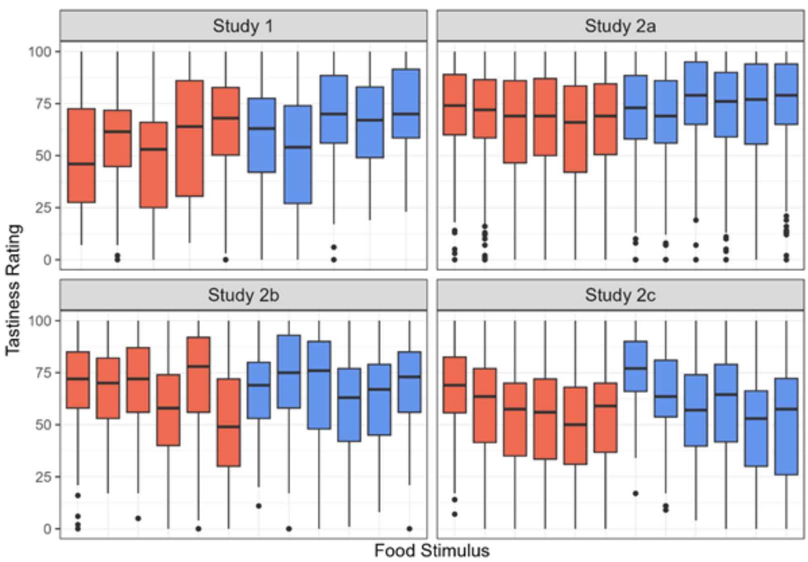
Discussion of Studies 2a–2c

The objective of Studies 2a–2c was to test the replicability of the erasure effect on subjective tastiness ratings initially documented in Study 1. Each of these three studies employed a newly drawn, independently and randomly selected set of 12 food images (six per condition), while maintaining the same within-subjects design, the same dependent variable (mean tastiness rating), and the same Bayesian sequential analysis protocol. The only design modification relative to Study 1 was the use of an improved erasure protocol, the display of 12 instead of 10 food images, slightly modified sampling and shuffling procedures, and the presentation of all trials within a single session rather than across 10 separate sessions.

In all three studies, the results consistently supported H1: participants rated food images as more appetizing when the objective food parameters (calorie values and stimulus IDs) had been stored (non-erasure condition) compared to when they had been erased (see Fig. 1). The Bayesian evidence was strong to very strong in each study (Study 2a: BF_{10} = 26 × 1015, d = .50; Study 2b: BF_{10} = 27.23, d = .29; Study 2c: BF_{10} = 10.40, d = .26). Together with the result obtained in Study 1 (BF_{10} = 1309.22, d = .66), the erasure effect on subjective tastiness ratings was thus documented across four independent studies, each of which used a different set of randomly selected food images. This pattern of findings closely mirrors the results reported by Maier and Dechamps (2025), who observed a consistent erasure effect on subjective likability ratings across four studies with different sets of randomly selected color stimuli.

Figure 1

Mean Ratings of Food Stimuli in the Erasure (Red) and Non-Erasure (Blue) Conditions for Each Study



The convergence between the present food-erasure results and the original color-erasure findings is noteworthy. In both paradigms, subjective evaluations of color likability or food tastiness were systematically higher in the non-erasure condition. The effect was obtained with entirely different stimulus domains (colors vs. food images), different objective parameters (hue and lightness vs. calorie values and stimulus IDs), and different subjective assessment dimensions (likability vs. tastiness). This cross-domain consistency strengthens the interpretation that the observed effects may reflect a general phenomenon of macroscopic complementarity between subjective evaluations and objective measurements, as proposed by the non-commutability conjecture of the GQT.

However, as was the case in each of the individual color-erasure studies reported by Maier and Dechamps (2025), the interpretation of the present findings remains ambiguous at the level of each individual study. Because each study used a limited set of food images fixedly assigned to conditions, the alternative explanation, i.e. that the specific stimulus sets coincidentally favored H1, cannot be ruled out within any single study. Although the probability of such a favorable assignment occurring by chance in four consecutive studies with four independently selected stimulus sets becomes increasingly small, this alternative explanation remains formally viable when each study is considered in isolation.

Our research agenda involved three separate steps. First, the erasure effect needed to be documented in a first study (Step 1). This goal was achieved in Study 1, as reported above. Second, the effect needed to be shown to be replicable across several independent studies with different stimulus material (Step 2). This was also accomplished, as Studies 2a–2c all yielded strong Bayesian evidence for the predicted effect with three different sets of food images. Third, the alternative factor of a biased stimulus assignment argument needs to be ruled out by a final overall analysis (Step 3) to rectify the internal validity violation within the experimental designs, thereby allowing a clear attribution of the effects found so far to the erasure manipulation. This three-step research agenda will next be completed by reporting the results of the preregistered overall analyses.

Overall Analyses for Tastiness: Testing the Impact of Biased Stimulus Assignments

Across the four studies, the experimental manipulation produced a consistent effect on the primary dependent variable, participants' subjective tastiness ratings of food images. This pattern is analogous to the findings reported for the color-based paradigm by Maier and Dechamps (2025) and appeared robust across all four independent samples. However, internal validity is limited because, in each study, the erasure manipulation was intentionally confounded with stimulus assignment: only a restricted subset of food pictures was randomly allocated to the erasure versus non-erasure condition (Study 1: 5 vs. 5; Studies 2a–2c: 6 vs. 6). As a result, the observed effect cannot be attributed unequivocally to the erasure manipulation; instead, it may reflect a biased assignment of specific food stimuli to experimental conditions. In Step 3 of our research agenda, we therefore subjected this alternative stimulus-assignment explanation to preregistered frequentist statistical tests, specifically a stratified permutation analysis and a multilevel model. To this end, data from all four studies were combined into a single data file (N_{total} = 647).

The stratified permutation analysis and all subsequent analyses reported here were preregistered after data collection was completed. At the time of preregistration, the Bayesian t-test results for the individual studies were already known to the authors; however, neither the stratified permutation analysis nor the multilevel model analysis had been conducted. The preregistration can be accessed at OSF (https://osf.io/ghkr7). Note that in the preregistration the sample size provided for each study in the table was based on the total number of participants including exclusions ($N_{total—with—exclusions}$ = 656). The preregistered analyses reported here, however, were performed on the data without exclusions to be consistent with all data analyses of Steps 1 and 2.

The independent variable was the within-subject erasure manipulation, comprising a non-erasure condition in which objective calorie information and the image ID were stored in the results file, and an erasure condition in which these objective data were permanently deleted and therefore cannot be reconstructed retrospectively. The dependent variable was the subjective tastiness rating, with which participants evaluated each presented food image in terms of how appetizing or tasty the depicted food appeared.

Primary Analysis: Stratified Permutation Test

The purpose of the stratified permutation analysis was to evaluate the hypothesis (H1) that the observed overall erasure effect across the four studies cannot be attributed to a coincidentally favorable stimulus-to-condition assignment (one-sided). In other words, H1 states that the effect reflects a genuine manipulation effect rather than an artifact of biased stimulus allocation. To quantify this, we estimated, via stratified within-study permutations, the probability of obtaining an overall effect at least as large as the observed one if stimulus assignment to erasure condition were effectively random. A small permutation p-value ($\alpha = .05$) would therefore support H1 by indicating that a fortuitous stimulus allocation is unlikely to account for the observed effect.

The procedure permutes stimulus assignments to conditions independently within each study, thereby respecting the nested data structure (stimuli within studies) without assuming that stimuli are exchangeable across studies. For each study, the number of possible assignments of stimuli to conditions is fixed by the study design: Study 1 allows 252 assignments (5 vs. 5), and Studies 2a–2c each allow 924 assignments (6 vs. 6). The full assignment space is the Cartesian product of the study-level assignment sets (approximately 199 million possible combinations). Because evaluating all combinations would be computationally burdensome, we used Monte Carlo sampling and drew 1,000,000 random combinations from this space.

The experimental procedure involved two random components: (a) the selection of 10 (Study 1) or 12 (Studies 2a–2c) stimuli from a larger pool of 100 food images, and (b) the assignment of the selected stimuli to the two experimental conditions. The stratified permutation analysis addresses component (b) by sampling from the full set of possible within-study assignments. Component (a) cannot be tested via permutation because subjective ratings are available only for the stimuli that were actually selected and shown. However, component (a) primarily concerns generalizability rather than confounding. It is addressed by the replication across four independent studies, each using an independently drawn stimulus set. The consistency of the erasure effect across these independently drawn sets, formally evaluated by the stratified permutation analysis, provides evidence that the overall result is not driven by any single fortunate stimulus draw.

First, we computed stimulus-level mean tastiness ratings separately for each study and condition. For each study, we then quantified the study-specific effect as the difference between the mean of the non-erasure stimulus means and the mean of the erasure stimulus means. Because the design was within-subjects and all participants rated all stimuli within a study, stimulus means were based on the same participant sample within each study, so no additional within-study weighting was applied. The condition means (non-erasure vs. erasure) and corresponding differences (non-erasure minus erasure) were as follows: Study 1, $M = 63.2$ vs. 55.1 (difference = 8.14; $n = 55$; 5 vs. 5 stimuli); Study 2a, $M = 72.1$ vs. 67.0 (difference = 5.10; $n = 363$; 6 vs. 6 stimuli); Study 2b, $M = 66.2$ vs. 63.4 (difference = 2.79; $n = 121$; 6 vs. 6 stimuli); and Study 2c, $M = 60.0$ vs. 56.2 (difference = 3.76; $n = 108$; 6 vs. 6 stimuli). These study-specific differences were then aggregated into a single overall test statistic by computing a weighted mean across studies, using each study's participant sample size (n) as the weight, so that larger studies contributed more to the overall estimate. This yielded an observed overall mean difference of 4.70 points (non-erasure minus erasure), which reflects the sample-size-weighted average effect across all four studies.

To assess whether this overall effect could be explained by coincidental stimulus-to-condition assignment, we conducted the stratified permutation test in which stimulus-level mean ratings were randomly reassigned to conditions within each study while keeping the original number of stimuli per condition fixed. In other words, we permuted the assignment of stimulus means to conditions (rather than participant labels), thereby isolating the potential impact of chance stimulus allocation. For each permutation, we recomputed the study-specific mean differences (non-erasure minus erasure) and then combined them into the overall test statistic using the exact same computation as for the observed effect, namely a sample-size–weighted mean across studies. Based on 1,000,000 Monte Carlo samples from the full assignment space, the probability of obtaining an overall difference at least as large as the observed one in the predicted direction (non-erasure > erasure) was p = .002 (Monte Carlo SE = .000046), corresponding to 2,148 out of 1,000,000 sampled assignments. Thus, the observed overall effect is unlikely to be attributable to a fortuitous stimulus allocation.

Figure 2

Distribution of the Weighted Mean Difference Across 1,000,000 Stratified (Within-Study) Permutations (Observed Difference Indicated in Red)

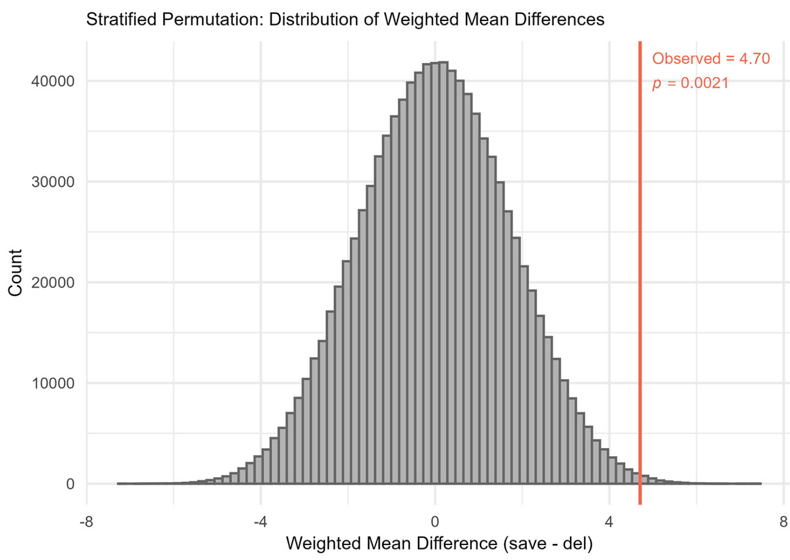


Figure 2 illustrates the permutation null distribution of the weighted overall mean differences and shows that the observed effect falls in the extreme right tail of the distribution.[1]

1 It can be argued that while there is no issue in expanding Bayesian evidence for individual studies

Secondary Analysis: Multilevel Model

Furthermore, we tested whether the erasure effect on subjective tastiness ratings remains significant when controlling for stimulus-specific variance using a multilevel model (MLM; H2, two-sided: mean tastiness ratings differ between non-erasure and erasure). Subjective tastiness ratings served as the DV, and condition (non-erasure vs. erasure) was specified as the IV (fixed effect). To account for stimulus-specific variance, we included random intercepts for stimuli nested within studies. In addition, we included study as a random intercept to model cross-study heterogeneity and to capture design differences between Study 1 and Studies 2a–2c.

The MLM yielded a significant fixed effect of condition, indicating that tastiness ratings were higher in the non-erasure condition than in the erasure condition, $\beta = 4.72$ ($SE = 1.76$), $t(39.11) = 2.68$, $p = .011$, corresponding to a small effect ($d = .19$).

Variance decomposition showed that a small proportion of variance was attributable to stimuli nested within studies (Var = 30.53; $ICC_{stimulu}s = .047$) and a similarly small proportion to between-study heterogeneity (Var = 25.54; $ICC_{study} = .039$), whereas most variance remained at the residual level (Var = 590.62). Thus, although some systematic differences between stimuli (and studies) were present, the bulk of variability reflected individual responses rather than stimulus-driven or study-driven effects.

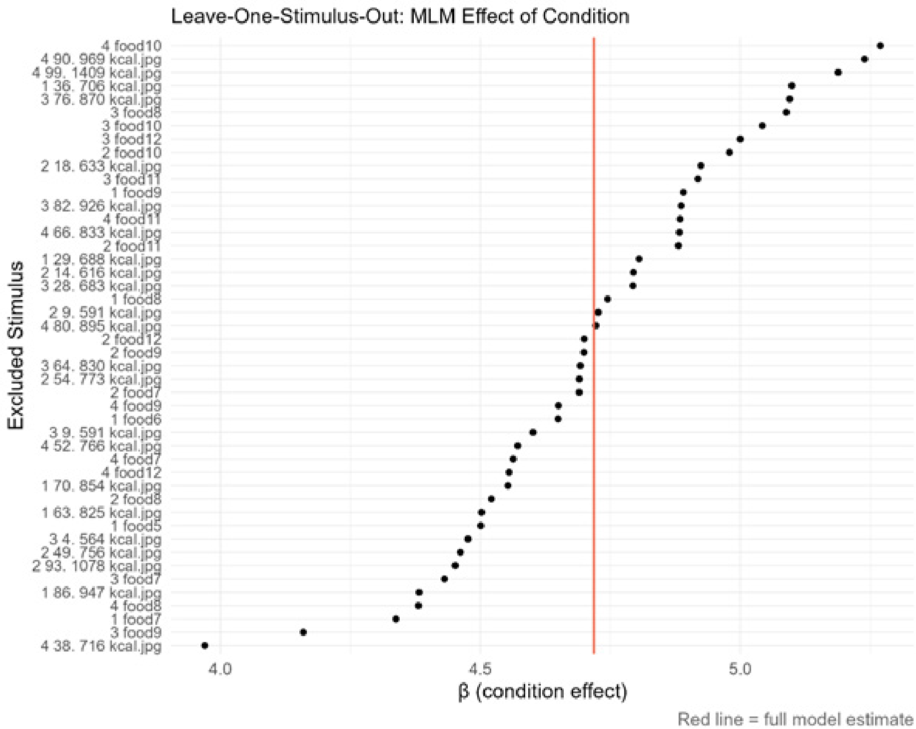
Sensitivity Analyses

Leave-one-stimulus-out Analysis. A leave-one-stimulus-out sensitivity analysis confirmed that the condition effect was stable. Specifically, we refit the MLM while excluding each stimulus in turn and tracked the resulting fixed effect estimate for condition. As shown in Figure 3, the estimated effects clustered closely around the full-model estimate (dashed red line at $\beta = 4.72$) and varied only minimally (β range = 3.97–5.27; all *p*s < .05). No single stimulus exclusion substantially altered the effect or its significance.

beyond the a priori specified evidence criterion, in case of aggregated data there can be an overrepresentation of favorable data, which might bias a permutation test. We therefore trimmed the experimental data to only include data from each study up until the first time $BF \geq 10$ (for studies 2a to c also with a pre-specified minimum $n > 100$) was met and submitted them to a further permutation analysis. The resulting sample consisted of $n = 355$ participants and showed a weighted mean difference $d = 4.67$ between conditions. Based on 1,000,000 Monte Carlo samples from the full assignment space, the probability of obtaining an overall difference at least as large as the observed one in the predicted direction (non-erasure > erasure) was $p = .008$ (Monte Carlo $SE = .000087$).

Figure 3

Leave-One-Stimulus-Out Sensitivity Analysis of the MLM Condition Effect (Erasure vs. Non-Erasure). Points Show β Estimates; the Red Line Shows the Full-Model Estimate (β = 4.72).



Study-Level Permutation p-Values and Combination Methods. To further assess robustness across studies, we computed exact (exhaustive) within-study permutation p-values for each study separately. The observed differences (non-erasure minus erasure) were positive in all four studies (Study 1: 8.14, $p = .040$; Study 2a: 5.10, $p = .011$; Study 2b: 2.79, $p = .235$; Study 2c: 3.76, $p = .234$), indicating consistent directionality but reduced statistical power in Studies 2b and 2c when considered in isolation. Combining the four study-level *p*-values yielded significant overall evidence across two standard combination methods (Edgington: p = .003; Fisher: p = .006), converging with the preregistered stratified permutation result and supporting the conclusion that the overall effect is unlikely to be explained by fortuitous stimulus-to-condition assignment.

Taken together, these findings suggest that the erasure manipulation effect on tastiness ratings is robust and not driven by specific stimuli.

Discussion

In the overall analyses provided here, Step 3 of the stepwise research agenda was performed. These analyses addressed the problem of a non-ambiguous causal interpretation of the experimental erasure manipulation on tastiness ratings by ruling out the main alternative explanation of "lucky" random stimulus assignments to experimental conditions (stratified permutation analysis and supplementary MLM).

The preregistered stratified permutation analysis provided the most direct test of our central concern. This analysis indicated that the likelihood of finding such an overall experimental erasure effect across studies by chance assignment of stimuli to conditions was very low (p = .002). By permuting stimulus assignments to conditions independently within each study and computing a sample-size-weighted overall test statistic across 1,000,000 Monte Carlo samples, this analysis directly addressed the "lucky assignment" alternative explanation while respecting the nested data structure and within-subjects design of the individual studies. In contrast to the pooled permutation approach used by Maier and Dechamps (2025) for the color-erasure data, the stratified approach employed here did not assume exchangeability of stimuli across studies, which is a methodological refinement that more accurately reflects the independent stimulus selection in each study.

The supplementary MLM analysis further supported the robustness of the experimental erasure effect. The results showed that the significant effect of the erasure manipulation (β = 4.72, p = .011) persisted after controlling for stimulus-specific variation within the combined sample. The low ICC values for stimuli (.047) and studies (.039) indicated that most variance came from individual responses rather than systematic stimulus or study effects. The leave-one-stimulus-out sensitivity analysis confirmed that the condition effect remained stable and significant when any single stimulus was excluded, demonstrating that the overall effect was not driven by any particular food image. The sensitivity analyses using alternative p-value combination methods (Edgington: p = .003; Fisher: p = .006) converged with the stratified permutation result.

In sum, the stratified permutation analysis provided compelling evidence that the observed effects were not due to chance ("lucky") stimulus assignments, while the MLM analysis demonstrated robustness within the observed stimulus set and ruled out study-level confounds. Together, these findings support the conclusion that the erasure manipulation of objective food parameters itself had an effect on subjective tastiness ratings. These analyses rectified the internal validity violation within the experimental settings, making an overall causal interpretation in a retrospective manner possible. They constitute the final Step 3 of our research agenda, which addressed the problem of testing an acausal phenomenon with a scientific-causal approach. Steps 1 and 2 tested the existence of an experimental effect and its replicability across several studies, in each of which an objective data-erasure interpretation on subjective data was confounded with a "biased stimulus assignment to experimental conditions" interpretation. This confound was designed to stabilize the assumed acausal subjective-objective non-commutability relation even under experimental-causal testing conditions within studies. In a third step, the confound was disentangled through overall analyses of the combined dataset. This procedure followed similar recommendations made by Römer (2023b) and Lucadou et al. (2007) and closely parallels the strategy successfully employed by Maier and Dechamps (2025) in the color-erasure paradigm.

As a consequence, these findings lend support to the conjecture derived from the GQT that macroscopic complementary relations exist between objective and subjective measurements of food stimuli. The results indicate that food images are perceived as more appetizing when the objective food parameters are retained compared to a condition in which they are erased. Step 3 analyses thus provided further evidence for the non-commutability conjecture derived from the GQT (Atmanspacher et al., 2002; Walach & Römer, 2000, 2011; see also Fach, 2011; Filk & Römer, 2011; Hinterberger & von Stillfried, 2013; Römer, 2023a, 2023b; Lucadou et al., 2007; Walach & von Stillfried, 2011), which proposed in the present experiments that experimental variations in objective food parameter documentation bias subjective experience of the food stimuli under investigation. Critically, this evidence was obtained in a stimulus domain entirely different from the original color-erasure studies, demonstrating that the non-commutability phenomenon generalizes beyond color perception.

General Discussion

The present research sought to conceptually replicate the macroscopic complementary relations between subjective evaluations and objective measurements documented in the color domain by Maier and Dechamps (2025) using food images as an entirely different class of stimuli. Following the non-commutability conjecture of the GQT, we predicted that erasure of objective food parameters (calorie values and stimulus IDs) would affect subjective tastiness ratings. The three-step research agenda originally developed by Maier and Dechamps (2025) was applied here with the same rationale regarding the deliberate confound and retrospective causality testing and successfully completed across all three steps.

In Steps 1 and 2, the erasure effect on subjective tastiness ratings was detected in Study 1 and successfully replicated across Studies 2a–2c. In all four studies, participants rated food images as more appetizing in the non-erasure condition than in the erasure condition, with strong to very strong Bayesian evidence (Study 1: $BF_{10} = 1309.22$, $d = .66$; Study 2a: $BF_{10} = 26 \times 10^{15}$, $d = .50$; Study 2b: $BF_{10} = 27.23$, $d = .29$; Study 2c: $BF_{10} = 10.40$, $d = .26$). In Step 3, the preregistered overall analyses – stratified permutation ($p = .002$), multilevel model ($\beta = 4.72$, $p = .011$), and sensitivity analyses (Edgington: $p = .003$; Fisher: $p = .006$) – ruled out the biased stimulus assignment alternative, leaving the erasure manipulation as the most probable explanation.

The central contribution of the present research lies in the cross-domain generalizability of the erasure effect. Whereas Maier and Dechamps (2025) documented non-commutability between subjective likability and objective color parameters (hue and lightness), the present studies demonstrate the same phenomenon with entirely different stimulus material (food images), different objective parameters (calorie values and stimulus IDs), and a different subjective assessment dimension (tastiness). The consistency of the erasure effect across these two independent stimulus domains substantially strengthens the evidence for the non-commutability conjecture of the GQT. It suggests that the phenomenon is not tied to the specific characteristics of color perception but may reflect a more general property of the psychophysical system – one in which the objective documentation of stimulus parameters systematically co-relates with subjective evaluative judgments.

With respect to the testing strategy, the same considerations raised by Maier and Dechamps (2025) apply. The retrospective overall analyses re-established a causal interpretation of the data without constituting a strict causality test at the study level. We consider this distinction critical given the NT axiom's prohibition of robust causal utilization of macroscopic entanglement correlations (Römer, 2023b; Lucadou et al., 2007). This indirect approach to causality testing, discussed in detail by Maier and Dechamps (2025; see also Walach et al., 2022), appears to circumvent the scientific causality-testing paradox inherent in the GQT. Importantly, the present findings complement the strict causality test conducted by Maier et al. (2026), who demonstrated the predicted effect-and-decline pattern in an unconfounded color-erasure study. Together, confound-protected stability (present research and Maier & Dechamps, 2025) and

NT axiom-driven decline under strict testing (Maier et al., 2026) form a coherent empirical picture fully consistent with the theoretical predictions of the GQT.

The non-local nature of the documented interactions is supported by the fact that participants were unaware of the erasure manipulation and that the objective data were stored or deleted after subjective responses had been recorded, precluding classical-local information transfer. The robustness and replicability of the effect across four studies with independently drawn stimulus sets lends support to the reclassification of these erasure-dependent effects as induced correlations (see the introduction), consistent with the Pauli-Jung mind-matter framework (Atmanspacher & Rickles, 2022). As in the color domain (Maier & Dechamps, 2025), the acausal entanglement correlations underlying the food-erasure effects were not limited by the NT axiom because strict causality tests were avoided at the study level.

Maier and Dechamps (2025) discussed the implications of the color-erasure findings for central positions in the philosophy of mind, arguing that the data challenge both Descartes' substance dualism (Descartes, 1641) and physicalism (Stoljar, 2024), while aligning with dual-aspect monism (Atmanspacher & Rickles, 2022). The present cross-domain replication reinforces these conclusions. The fact that macroscopic complementary relations were now documented in two independent stimulus domains makes it increasingly difficult to dismiss the findings as domain-specific artifacts. If subjective and objective aspects of reality were either ontically separated (substance dualism) or reducible to a single physical substrate (physicalism), the systematic co-relation between erasure of objective parameters and changes in subjective evaluations across colors and food images alike would remain unexplained. Only if subjective experience constitutes a non-physical yet complementary aspect of reality, as proposed by the GQT and dual-aspect monism, can macroscopic complementarity phenomena of the kind documented here be expected to occur (de Broglie, 1958; Römer, 2023b; Zeh, 2012).

Taking these findings together with the convergent evidence from the color-erasure studies (Maier & Dechamps, 2025) and the strict causality test (Maier et al., 2026), a speculative yet coherent picture of the nature of psychophysical reality begins to emerge. After measurement, subjective experience and objective physical reality appear separated and present themselves to the observer on an epistemic level as two distinct, complementary perspectives on the world. However, on an ontological level, both aspects may be fundamentally unified: prior to measurement, they exist as an undifferentiated, acausally entangled whole. The act of scientifically documenting psychophysical relations with a strict causal testing protocol – including direct replications, the gold standard of documenting objectively existing phenomena in the natural sciences – transfers this acausal, entangled reality into a classical, spatiotemporally local, and causally structured psychophysical reality by effectively destroying its true acausal nature. In this view, causality and spatiotemporal locality are epistemological features that emerge through the act of scientific documentation, whereas the underlying ontological psychophysical reality is acausally connected, complementary, and non-locally entangled. The deliberate use of a confounded experimental design, which implied a reduced objective testing strategy, in the present research may have preserved this acausally connected structure despite the act of phenomenon documentation, allowing the complementary relation to remain intact and the erasure effect to be stably documented. In other words, the present research may have shifted the existing metaphysical boundary, which presently demarcates empirically objectifiable from non-objectifiable phenomena, toward a new boundary that extends the reach of empirical documentation into a domain previously considered non-objectifiable, albeit at the cost of reduced objectifiability.

If we understand experience-based perception and volitional influence as the prototypical manifestations of induced correlations within this framework, then their seemingly unidirectional causal descriptions on the epistemic level – from objective to subjective (perception) or from subjective to objective (volition) – may in fact represent complementary perspectives on an ontologically singular, acausal psychophysical process. What we subjectively prefer, we tend to manifest more readily in the objective, material world and conversely, what is objectively more fully realized tends to be subjectively perceived as more appealing. Experience-based perception and volitional influence would therefore not constitute two separate causal mechanisms but rather two complementary descriptions of the same underlying acausal psychophysical phenomenon, differing only in the direction of the causal narrative imposed upon them. This speculative account, while requiring further theoretical and empirical elaboration, offers a coherent integrative framework for understanding the acausal complementary relations documented across the present and previous studies within the broader context of the GQT and dual-aspect monism.

It should be noted that the results presented here do not provide an exhaustive test of this theoretical framework. While the empirical evidence for the GQT is still accumulating, the cross-domain replication achieved in the present research together with the convergent evidence from the color-erasure studies (Maier & Dechamps, 2025; Maier et al., 2026) suggests a promising outlook for its future empirical evaluation.

Acknowledgments

We would like to thank the students of the two Empra Courses for their support during data collections of Studies 1 and 2b and for developing the design.

Finally, we would like to thank several bachelor students for their support in data collection of Studies 2a and 2c.

Authors' Contributions

MAM proposed the theory and originally co-designed the studies. He drafted the first version of the manuscript. AV co-designed the studies, programmed the studies, helped in data preparation and was actively involved in revising the manuscript. MCD co-designed the studies, programmed the studies, helped in data preparation and was actively involved in revising the manuscript.

Declaration of Interests

The Authors declare that there is no conflict of interest.

References

Atmanspacher, H., & Rickles, D. (2022). *Dual-aspect monism and the deep structure of meaning*. Routledge. https://doi.org/10.4324/9781003270584

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

Bohr, N. (1928). The quantum postulate and the recent development of atomic theory. *Nature, 121*, 580–590. https://doi.org/10.1038/121580a0

Chalmers, D. J. (1995). Facing up to the problem of consciousness. *Journal of Consciousness Studies, 2*, 200–219.

Chalmers, D. J. (1996). *The conscious mind: In search of a fundamental theor*y. Oxford University Press.

de Broglie, L. (1958). *La théorie de la mesure en mécanique ondulatoir*e. Gauthier-Villars.

Descartes, R. (1641). *Meditationes de prima philosophia*. Parisiis apud Michaelem Soly.

Edgington, E. S. (1972). An additive method for combining probability values from independent experiments. *The Journal of Psychology, 80*(2), 351–363.

Fach, W. (2011). Phenomenological aspects of complementarity and entanglement in exceptional human experiences (ExE). *Axiomathes, 21*, 233–247. https://doi.org/10.1007/s10516-010-9143-7

Filk, T., & Römer, H. (2011). Generalised quantum theory: Overview and latest developments. *Axiomathes, 21*(2), 211–220. https://doi.org/10.1007/s10516-010-9136-6

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

Hinterberger, T., & von Stillfried, N. (2013). The concept of complementarity and its role in quantum entanglement and generalized entanglement. *Global Philosophy, 23*, 443–459. https://doi.org/10.1007/s10516-012-9187-y

Levine, J. (1983). Materialism and qualia: The explanatory gap. *Pacific Philosophical Quarterly, 64*, 354–361. https://doi.org/10.1111/j.1468-0114.1983.tb00207.x

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 color. *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), 16–48. https://doi.org/10.31156/jaex.23535

Maier, M. A., Vogel, A., Storch, J., & 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

Nagel, T. (1974). What is it like to be a bat? *The Philosophical Review, 83*, 435–450. https://doi.org/10.2307/2183914

Römer, H. (2023a). Twenty years of Generalized Quantum Theory. *Journal of Anomalistics, 23*, 145–154.

Römer, H. (2023b). *Quanten, Komplementarität und Verschränkung in der Lebenswelt. Verallgemeinerte Quantentheorie* (Perspektiven der Anomalistik, Bd. 7). Lit Verlag.

Ruffing, E. (2021). *Einführung in die Geschichte der Philosophie* (3rd ed.). Brill Schöningh. https://doi.org/10.36198/9783838557595

Shariff, A. F., Schooler, J., & Vohs, K. D. (2008). The hazards of claiming to have solved the hard problem of free will. In J. Baer, J. C. Kaufman, & R. F. Baumeister (Eds.), *Are we free? Psychology and free will* (pp. 181–204). Oxford University Press. https://doi.org/10.1093/acprof:oso/9780195189636.003.0009

Stoljar, D. (2024). Physicalism. In E. N. Zalta & U. Nodelman (Eds.), *The Stanford encyclopedia of philosophy* (Winter 2024 ed.). Stanford University. https://plato.stanford.edu/archives/win2024/entries/physicalism/

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

Walach, H., & Römer, H. (2011). Generalized entanglement – A nonreductive option for a phenomenologically dualist and ontologically monist view of consciousness. In H. Walach, S. Schmidt, & W. B. Jonas (Eds.), *Neuroscience, consciousness and spirituality* (Vol. 1, pp. 81–95). Springer. https://doi.org/10.1007/978-94-007-2079-4_6

Walach, H., & von Stillfried, N. (2011). Generalized Quantum Theory-Basic idea and general intuition: A background story and overview. *Axiomathes, 21*(2), 185–209. https://doi.org/10.1007/s10516-010-9145-5

Walach, H., Horan, M., Hinterberger, T., & Lucadou, W. v. (2022). Evidence for anomalistic correlations between human behavior and a random event generator: Result of an independent replication of a micro-PK experiment. *Psychology of Consciousness: Theory, Research, and Practice, 9*(2), 173–188. https://doi.org/10.1037/cns0000233

Zeh, H. D. (2012). *The physical basis of the direction of time* (5th ed.). Springer. https://doi.org/10.1007/978-3-540-68001-7

Zusammenfassung

Makroskopische Komplementarität zwischen subjektiven und objektiven Bewertungen von Lebensmittelbildern: Eine konzeptionelle Replikation des „Color-Erasure"-Effekts unter Berücksichtigung von Störeinflüssen

Die Verallgemeinerte Quantentheorie (GQT) postuliert, dass subjektive Erfahrungen und objektive physikalische Messungen makroskopische komplementäre Teilsysteme bilden, die durch akausale, nicht-lokale Verschränkungskorrelationen miteinander verbunden sind. Die Nicht-Kommutativitätsannahme der GQT sagt voraus, dass eine Messung innerhalb eines Teilsystems den Zustand des Gesamtsystems verändert, einschließlich der Eigenwerte des komplementären Teilsystems. Maier und Dechamps (2025) lieferten erste Evidenz für diese Annahme, indem sie zeigten, dass die Speicherung (Nicht-Löschung) versus Löschung objektiver Farbparameter (Farbton und Helligkeit) systematisch die subjektiven Gefallensbewertungen über vier Studien hinweg beeinflusste (höherer subjektiver Gefallens-Mittelwert in der Nicht-Löschungsbedingung im Vergleich zur Löschungsbedingung). Die vorliegende Forschung replizierte diese Farb-Löschungsbefunde konzeptuell in einem anderen Stimulusbereich. In vier Studien (Gesamt-N = 647) bewerteten die Teilnehmenden Fotografien von Lebensmittelgerichten hinsichtlich ihrer subjektiven Schmackhaftigkeit, während die objektiven Lebensmittelparameter (Kalorienwerte und Stimulus-IDs) entweder gespeichert (Nicht-Löschungsbedingung) oder dauerhaft gelöscht wurden (Löschungsbedingung). Wie in den Farbstudien wurden begrenzte

Stimulus-Sets den Bedingungen innerhalb jeder Studie fest zugeordnet, wodurch bewusst ein Stimulus-Bias-Konfund eingeführt wurde, um die akausale Komplementärrelation vor der zerstörerischen Wirkung des Nicht-Transmissions-Axioms (NT-Axiom) zu schützen. Gemäß einem vordefinierten dreistufigen Analyseplan dokumentierte die präregistrierte Studie 1 (Schritt 1) einen starken Löschungseffekt auf die Schmackhaftigkeitsbewertungen. Die Teilnehmenden bewerteten Lebensmittelgerichte in der Nicht-Löschungsbedingung als appetitlicher im Vergleich zur Löschungsbedingung (BF_{10} = 1309,22, d = .66), und die Studien 2a–2c (Schritt 2) replizierten diesen Effekt mit drei unabhängig gezogenen Stimulus-Sets (BF_{10}-Werte von 10,40 bis 26×10^{15}, d-Werte von .26 bis .50). In Schritt 3 schlossen präregistrierte Gesamtanalysen – einschließlich eines stratifizierten Permutationstests (p = .002), eines Mehrebenenmodells (β = 4,72, p = .011) und Sensitivitätsanalysen – die alternative Erklärung einer verzerrten Stimuluszuordnung aus. Diese Befunde liefern domänenübergreifende Evidenz für die Nicht-Kommutativitätsannahme: Die Dokumentation objektiver Stimulusparameter scheint akausal mit subjektiven Bewertungen zu korrelieren, konsistent mit makroskopischen Komplementärrelationen, wie sie von der GQT vorhergesagt werden. Die Implikationen für den Substanzdualismus, den Physikalismus und den Duale Aspekte Monismus werden diskutiert.

Schlüsselbegriffe: Verallgemeinerte Quantentheorie, makroskopische Komplementarität, makroskopische non-lokale Verschränkungskorrelation, Nicht-Kommutativität, Löschungsparadigma, psychophysische Interaktion