Journal of Anomalistics Volume 24 (2024), pp. 15–54

Complementarity of Subjective and Objective Realities

An Experimental Investigation of Variations of Subjective Self-reported Intelligence Data when Objective Data are Erased or Not-erased

Markus A. Maier, Chantale Geissler, Emilia C. DeMattia, Anastasia Vogel, Moritz C. Dechamps[1]

Abstract – The Complementarity Principle (CP), introduced by Nils Bohr, described the wave-particle duality of quantum phenomena and its dependence on the measurement setup exploring the quantum states. In general, this principle summarized the fact that two mutually exclusive and contradictory states of quantum events can be reconciled as a measurement-dependent occurrence of these states. Originally, the CP was postulated only for the realm of quantum mechanics, but recently the Generalized Quantum Theory (GQT) extended its area of applicability to the macroscopic domain. According to GQT complementarity relations are also hypothesized to exist between macroscopic objective and subjective measurements of the same concept. This implies objective measurement-dependent variations of subjective experience. To test this hypothesis, seven studies were conducted in which individual variations in subjective intelligence ratings were tested in relation to the availability or non-availability of corresponding objective individual intelligence data. Only one pre-registered study (Study 2) showed strong Bayesian evidence for H1 ($BF_{10} > 10$), indicating, as predicted, higher subjective intelligence ratings when objective data were erased compared to a condition in which objective data were available within a male sample. This effect could not be replicated in direct replication attempts, nor did a moderator search in subsequent studies find any robust systematic variation in the data. The results seem to question the validity of the macroscopic complementarity conjecture derived from the GQT. On the other hand, they could also be interpreted as the "effect and decline" data pattern that the Model of Pragmatic Information would predict when conducting empirical confirmations of macroscopic complementarity relations. Possible future research strategies to clarify these different interpretations are discussed.

Keywords: complementarity principle – Generalized Quantum Theory – erasure manipulations – subjectivity-objectivity duality – psi

https://doi.org/10.23793/zfa.2024.015

Komplementarität von subjektiven und objektiven Realitäten – Eine experimentelle Untersuchung der Variationen von subjektiven selbstberichteten Intelligenzdaten, wenn objektive Daten gelöscht oder nicht gelöscht werden

Zusammenfassung – Das von Nils Bohr eingeführte Komplementaritätsprinzip beschreibt den Welle-Teilchen-Dualismus von Quantenphänomenen und seine Abhängigkeit vom Messaufbau. Allgemein fasst dieses Prinzip die Tatsache zusammen, dass zwei sich ausschließende Quantenzustände unter Berücksichtigung der messbedingten Dokumentation als sich einander ergänzend betrachtet werden können. Ursprünglich fand das Komplementaritätsprinzip ausschließlich innerhalb der Quantenmechanik Anwendung, in jüngerer Zeit hat die verallgemeinerte Quantentheorie (Generalized Quantum Theory; GQT) dieses Prinzip auf den makroskopischen Bereich

---

1 **Markus Maier** has been Professor of General Psychology II (Emotion & Motivation) at LMU Munich since 2010. Prior to this, he was Assistant Professor of Psychology at the State University of New York at Stony Brook (USA) from 2008 to 2010. Since this time, he has been working on motivational psychology issues such as the question of free will and the control of behavior through intentions. He attempts to describe these theoretically on the basis of interactions between subjective and objective reality. As part of his empirical research, he and his working group are investigating an extension of Descartes' dualism that is necessary to describe free will.

Chantale Geissler has been studying psychology at LMU Munich since 2021 and philosophy since 2023. She is interested in the interface between psychology and philosophy and therefore also in the question of how the intentional influence on reality through motivational and emotional processes can exist despite Cartesian dualism. She has therefore been supporting the research of Markus Maier's working group since 2022.

Emilia DeMattia has been a bachelor student at LMU Munich since 2021. She first worked with Markus Maier in 2022 as part of an empirical internship and has been supporting the working group's research ever since. In addition to a special interest in clinical psychology, she is interested in motivational psychology and in particular in the question of the emergence of volitional impulses in the sense of free will. Maier's approach of investigating the interaction between subjective and objective reality in the context of multifaceted studies and in connection with various psychological constructs broadens the perspective and research-theoretical horizon of the mechanisms of the human psyche.

Anastasia Vogel completed her Bachelor of Science in Psychology at LMU Munich in 2023. During her Bachelor's thesis, she dealt with "subjective volition". She is currently working towards her Master's degree in Psychology, which she will complete in 2025. In addition, she is continuing the topic of free will and how objective reality can be created through subjective volition with Prof. Dr. Markus Maier. She is part of Prof. Dr. Markus Maier's working group, which aims to expand Decartes' dualism by describing free will.

Moritz Dechamps has been working as a research assistant at LMU Munich since 2015. He received his doctorate in 2019 with a dissertation on mind-matter interactions and their replicability and has since been researching the topics of consciousness, free will, anomalous cognition, expectation effects and self-fulfilling prophecies. He focuses in particular on the replicability and methodological approximation of non-linear effects and the use of technological methods.

angewendet. In der GQT werden Komplementaritätsbeziehungen auch zwischen objektiven und subjektiven Messungen eines makroskopischen Konzepts angenommen. Dies impliziert objektivierungsabhängige Veränderungen des subjektiven Erlebens. Um diese Hypothese zu überprüfen, wurden sieben Studien durchgeführt, in denen individuelle Variationen in subjektiven Intelligenzbeurteilungen in Abhängigkeit von der Verfügbarkeit bzw. Nichtverfügbarkeit entsprechender objektiver individueller Intelligenzdaten getestet wurden. Nur eine präregistrierte Studie (Studie 2) zeigte eine starke Bayes'sche Evidenz für die H1 (BF10 > 10). Für eine männliche Stichprobe kam es wie vorhergesagt zu einer höheren subjektiven Intelligenzbewertung, wenn objektive Daten gelöscht wurden, verglichen mit einer Bedingung, in der objektive Daten verfügbar waren. Dieser Effekt konnte in direkten Replikationsversuchen nicht repliziert werden, und die Suche nach Moderatoren in nachfolgenden Studien zeigte keine robuste systematische Variation in den Daten. Die Ergebnisse scheinen die Gültigkeit der aus der GQT abgeleiteten makroskopischen Komplementaritätshypothese in Frage zu stellen. Auf der anderen Seite könnten sie auch als das „Effekt und Decline"-Muster interpretiert werden, das das Modell der Pragmatischen Information bei der empirischen Bestätigung makroskopischer Komplementaritätsbeziehungen vorhersagen würde. Mögliche zukünftige Forschungsstrategien zur Klärung dieser unterschiedlichen Interpretationen werden diskutiert.

Schlüsselbegriffe: Komplementaritätsprinzip – Verallgemeinerte Quantentheorie – Löschmanipulationen – Subjektivitäts-Objektivitäts-Dualismus – Psi

Introduction

Niels Bohr introduced the *complementarity principle* at a physics congress in 1927, aiming to reconcile seemingly contradictory descriptions of light phenomena (Bohr, 1928, p. 580; Fahrenberg, 2013, pp. 300–301). Initially, this principle addressed the dual nature of light, which can be described both as a wave and as a particle. The complementary, i. e. phenomenologically contradictory, but mutually related description, results from the respective measuring arrangement, suggesting either a discrete (particle) or continuous (wave) state description of light, depending on the observation setting (Bohr, 1948, 1949). The complementarity principle is thus a direct consequence of the Copenhagen interpretation, which asserts that quantum physical outcomes cannot be defined independently of the measuring arrangement (see Favrholdt, 1999; pp. 24–25). While the complementarity principle was originally applied only within the realm of quantum physics due to the experimental demonstration of quantum-measurement dependence, Bohr speculated early on about its relevance to psychology (Fahrenberg, 2013; Favrholdt, 1999).

Bohr (1929) identified the measurement-dependent description of quantum states as central to complementarity. It quite naturally follows from the inherent impossibility of strictly separating phenomena from observation methods within quantum mechanics. This led him to contemplate whether a similar measurement dependency might exist in the study of certain psychological phenomena. For example, the relationship between subjective experience of certain psychological states and their objective documentation could reflect such a complementarity structure. Specifically, he mentioned the theoretical conflict between the feeling of free will and the seemingly "unbroken causal connection of the underlying physiological processes." Bohr argued that with the discovery of the quantum of action, a detailed causal tracing of atomic processes was not possible, since any attempt to determine them by measurement would lead to a basically uncontrollable intervention in their course. Bohr (1929, p. 486) concluded:

> According to the aforementioned view of the relation of brain processes and mental events, we must therefore be prepared for the fact that an attempt to observe the former would bring about a substantial change in the accompanying feeling of will. (Bohr, 1929, p. 486)

This assumption implies that any measurement-based attempt to objectify the material processes underlying behavior control, would lead to a marked change of the setting in which a subjective sense of free will had just occurred. Both, objective and subjective descriptions, can therefore not unequivocally be linked. The after-measurement setting is not the same as the before-measurement setting for which an experience of free will was originally reported. Thus, the existence of free will cannot be confirmed or rejected based on a measurement-dependent documentation of a deterministic behavior control through brain processes. Or in other words, free will situations and deterministic behavioral control situations cannot be determined by certain measurements simultaneously. This Gedankenexperiment nicely resembles the original wave-particle duality of light and the necessity of a complementary description of both aspects as suggested by Bohr (1929).

While modern quantum mechanics might dismiss the influence of measurements in macroscopic domains due to decoherence (Zeh, 1970), Bohr's early speculation highlights the potential parallels between subjective-objective relations in psychology and the wave-particle duality of light as complementary descriptions. Although Bohr (1929) acknowledged that his analogies might not hold precisely, he gingerly suggested a generalization of the complementarity principle to illuminate the subjectivity-objectivity relationships across various psychological domains.

Complementarity in Psychology

Psychology aims to elucidate the underlying mechanisms of human behavior. In motivational psychology, a distinction can be made between deterministic-causal approaches, in which behavior is determined by biological motive states and external incentives, meaning reward or punishment. Conversely, volitional approaches highlight consciously controlled and thus at least partially autonomous, intentional influence on behavior, with some theories integrating both perspectives (see Brandstätter et al., 2018; Rudolph, 2013).

Applying Bohr's complementarity principle to psychological explanations of behavioral control suggests an intriguing perspective. A deterministic-causal description of behavioral control can be seen as complementary to one emphasizing autonomous, free-will based control. Despite their apparent contradiction, these descriptions can be conceptualized as complementary due to the different measurement arrangements used to record them (see also Bohr, 1985).

Traditionally, complementarity has been seen as applicable only to quantum mechanical phenomena. However, recent theoretical developments propose a generalization of the complementarity principle to macroscopic domains (see Fahrenberg, 2013). In psychology, the Generalized Quantum Theory (GQT) of Walach and Römer (2000) and Atmanspacher et al. (2002; see also Fach, 2011; Filk & Römer, 2011; Römer, 2023; Lucadou et al., 2007; Walach & Stillfried, 2011) proposes complementarity relations between subjective-psychological and objective-physical concepts. In particular, the Model of Pragmatic Information (MPI; Lucadou et al., 2007) offers a framework for understanding the complementary interplay between the autonomous-subjective reality construction (i. e., willpower changes behavior and shapes reality in this way) and an objective-deterministic conception of psychological processes based on physicalist reality approaches (i. e., behavior and reality is shaped by deterministic neurocognitive processes). According to the MPI, both these forms of description are based on different degrees of pragmatic information involved in reality formation.

Pragmatic information, coined by Ernst von Weizsäcker (1974), denotes information intended to influence a receiver, comprising an effect-generating intention defined as novelty (N) and a knowledge structure shared between sender and receiver called confirmation (C). Both components are multiplicative and complementary and can be expressed by the formula: $I_{prag} = N \times C$. The MPI extends this concept to psychology, suggesting a complementary relationship between autonomy/novelty (uniqueness) of the subjective volitional impulse (= willpower) on the one hand and its objectifying confirmation on the other hand with both depending on the measurement arrangement applied (Lucadou, 1984, 1987, 1995, 1998, 2001, 2002, 2015).

Complementarity in this context means that the higher the impact of the objective component within the pragmatic information becomes, the lower the impact of the subjective-autonomous element (= intentional impulse) will be, and vice versa. Maier et al. (2022) applied this idea to volitional-autonomous reality construction in the context of volitional psychology. According to this theory, reality construction via conscious perception involves a process akin to measurement, wherein pragmatic information is conveyed. Objective, deterministic reality, as described by the natural sciences, involves passive registration excluding subjective volition.

Here, the confirmation factor dominates, and since it is maximum, any subjective impact is excluded (C = maximum → N = 0). In other words, if measurements are passive registrations and therefore constitute objective measurements any reality constructing intentional phenomena are automatically excluded and they cannot be empirically observed. As a consequence, they will be claimed non-existent. However, this is not an ontic inexistence but rather a measurement dependent one caused by the act of passive registration. So, to provide evidence for a volitionally constructed reality, which arises from an autonomous intentional impulse, one must reduce the confirmation element within the performed measurements. These reduced-objective measurements only allow for the assessment of novelty and thus subjective autonomous reality construction (C = reduced → N > 0). This understanding leads to a paradox in the free will debate: attempting to objectively document volitionally created realities undermines their subjective impact, mirroring Bohr's anticipation of measurement-dependent complementarity (Bohr, 1929).

To address this paradox, a new approach needed to be developed: A possible solution might lie at the core of the problem itself. Instead of trying the impossible (i. e. objectifying subjectively induced realities) one could empirically explore the assumed complementarity principle directly. Thus, in our research presented here, we tried to identify an empirically testable complementarity relation between a volitional reality creation and its objectification. Specifically, we propose that a subjectively willed reality about a psychological phenomenon will behave complementary depending on the experimentally manipulated presence or absence of objective data about the same phenomenon (see e. g., Maier et al., 2022). The corresponding empirically testable hypothesis states: A subjective reality constructed by volitional impulses varies with the degree of its objectification, that is, a more pronounced subjective bias along one's intention is predicted in a condition in which there is no or little subsequent objectification of this reality, and a less pronounced or no bias is expected in a condition in which there is subsequent objectification. In this way, a measurement-dependent description of subjectively created realities can be investigated. This was the aim of the studies reported below.

These studies focused on the psychological concept of "intelligence," which can be operationalized both by a subjective, introspective assessment that may be biased by the will of the self-reporting individual (see Alicke & Sedikides, 2009; Furnham, 2001; Neubauer & Hofer, 2020) and by an objective measure of intelligence that is largely free of subjective bias. Specifically, we investigated how subjective assessments of intelligence are influenced by subsequent objective intelligence documentation.

In a series of studies, participants provided subjective assessments of their intelligence before objective intelligence data was either erased (erasure condition) or stored (non-erasure condition). Importantly, participants were unaware of this erasure manipulation either at the time of self-report or later. The subjective evaluation of one's own intelligence should be biased by volitional impulses in the sense of pragmatic information ($E > 0$). Intelligence is widely considered to be highly relevant to motives (Alicke & Sedikides, 2009), suggesting an intentional bias. However, at the outset of this series of studies, it was unclear whether there was a normative tendency toward subjective overestimation or underestimation of self-report in the population studied. The volitional impulse underlying the subjective construction of reality in the context of intelligence assessment may be channeled in a particular direction by different socialization experiences, attitudes, and affectively charged beliefs, so that, in principle, both subjective overestimation and subjective underestimation of one's own intelligence might be expected. According to a study by Furnham (2001), participants on average estimated their IQ to be slightly higher than it actually is (see also Gignac & Zajenkowski, 2019; Neubauer & Hofer, 2020). Other studies (e. g., Engeler & Häubl, 2021) also reported lower subjective estimates of cognitive ability compared to objective ability levels. This heterogeneous pattern of findings highlights that there is an intentional bias in subjective assessments of intelligence and related concepts, but the direction of this effect may vary due to moderators. For this reason, the first study refrained from making an exact directional prediction and formulated an undirected hypothesis with a two-sided testing approach.

The specific hypothesis was that the subjective assessment of one's intelligence will differ depending on the treatment of objective data, reflecting the complementary relationship between subjective and objective perspectives. Participants remained uninformed about their objective data treatment throughout the whole experiment. The study was pre-registered online with the Open Science Framework (OSF) at https://osf.io/xty8a

Study 1

Methods

Ethical Guidelines

At the beginning of the survey, participants were generally informed about the study, the voluntary nature of their participation, and the data protection regulations (informed consent procedure). All stored data were encrypted and analyzed anonymously. Approval for the study was obtained from the responsible ethics committee of the Faculty of Psychology and Education at the LMU Munich.

Sample

A Bayesian sequential analysis procedure was used to analyze the data. This approach allowed for cumulative data collection and analysis, i. e., additional participants could be tested and new data could be successively added to the data set until a certain Bayes factor (BF) for H1 or H0 was reached. A priori, a stopping criterion of BF = 10 was set and pre-registered, i. e., data should be collected until a BF = 10 in favor of H0 or H1 was reached. If this was the case, data collection was stopped. If a BF > 10 was not reached within the previously defined maximum number of participants of $N = 3300$, data collection was also stopped (maximum N criterion).

Participants were recruited mainly through private contacts and social media by LMU students in the context of experimental courses under the supervision of Prof. Dr. Markus Maier.

In the end, 2,617 participants participated in the study. As pre-registered, participants were excluded from the data analysis if they answered more than two items of the objective intelligence test in less than five seconds (n = 65). Also excluded from analysis were individuals who indicated that they did not take the test in a quiet, undisturbed environment ($n = 364$) and/or who indicated that their responses were not reliable ("use-me-item," $n = 138$) or that they already knew the intelligence test used ($n = 45$). The last three statements were collected at the end of the study. The final sample included in the analysis consisted of $N = 2{,}109$ participants. The study was declared complete at this N, although the maximum N criterion or other stopping criterion had not been reached because resources for further data collection had been exhausted.

The sample consisted of 1,356 female, 740 male, and 12 diverse participants (1 record missing). The mean age of the participants was 32.30 years ($SD = 13.02$). Data collection was conducted online via PC, tablet, or smartphone, and participants could choose to perform the study in German or English language.

Materials

Subjective Intelligence Assessment

A 100-point visual analog scale was used to assess subjective intelligence, ranging from "not at all intelligent" (0) to "very intelligent" (100). Only the labels at the end of the scale were visible, not the numbers.

Objective Method for Intelligence Measurement

The six-item short version of the Hagen-Matritzen Test (HMT-S; Heydasch et al., 2013, 2017) was used to objectively assess nonverbal intelligence. A detailed description of this procedure can be found in the next section.

Procedure

Participants first selected their preferred language, English or German, and were informed of the study procedures. Subsequently, the participants agreed to participate (informed consent). A pseudorandom number generator then selected whether the participant was assigned to the erasure or non-erasure condition via jsPsych's *randomization*-module. This module relies on JavaScript's Math.random() function to generate pseudorandom numbers. Pseudorandom number generators are algorithms that produce sequences of numbers that mimic the properties of random numbers while being generated deterministically from an initial seed value. This method is commonly employed in experimental psychology to ensure the integrity and validity of research findings

It is important to note here that in this an all subsequently described experiments the participants were not aware of any objective data erasure or storage manipulations and they were never informed about to which condition they were assigned to nor that these conditions existed at all.

The next page was then presented, displaying the visual analog scale. Participants were asked to make a self-report of subjective intelligence by moving the slider to the appropriate position between the extremes of the scale. The final rating was then confirmed by pressing the button in the confirmation box. The objective intelligence test was then announced and explained. The tasks to be completed were explained and the time limit of a maximum of two minutes to complete each task was noted. Participants were then given the opportunity to familiarize themselves with the completion of the intelligence tasks on the following pages by means of two sample tasks, each of which was followed by the presentation and explanation of the correct answer. Then the six test items (matrices) were presented one after the other. No feedback was given this time. For each item, the participants were asked to identify a completion scheme and select the correct target item from five possible ones (single-choice format). The items were arranged in the form of a 3×3 group with 8 fields filled by matrices. The missing last field had to be filled in by selecting one of the 5 given options. The time to select a solution was a maximum of two minutes and was indicated by a countdown above the 9-field panel. After answering the item or after the maximum time had elapsed, the next item was automatically presented, in which case the previous item was scored as "incorrect." Participants' responses in the erasure condition were erased (overwritten with blank content) immediately after each item.

After completing the six test items, participants were asked to provide their demographic information (age and gender). Finally, three items asked whether the test had been administered in a quiet environment, whether the participants thought that their answers should (not) be included in the analysis (so-called "use-me" item), and whether they already knew the intelligence test.

The results were then stored in a result file containing either subjective data only (erasure condition) or subjective and objective data (non-erasure condition), depending on the condition.

Design and Statistical Analysis

The study design was a between-subjects design with an independent variable consisting of two conditions: the availability (non-erasure condition) or non-availability (erasure condition) of the objective IQ test scores in the final results file. The primary pre-registered statistical analysis consisted of a two-tailed Bayesian independent samples t-test with two independent groups (non-erasure vs. erasure) as the independent variable (IV) and subjective intelligence rating as the dependent variable (DV). As documented in the preregistration, we used an informed prior based on an estimated effect size of $d = .1$ following a Cauchy distribution centered around 0.05 with $r = .05$ (i.e., $\delta \sim$ Cauchy [.05, .05]). Data collection was conducted sequentially (by accumulating data in the order of participation) and analyzed in chronological order based on when participants completed the ratings.

Results

The mean subjective intelligence rating across all participants was 65.98 ($SD = 15.88$, range = 0 to 100), and the mean number of correctly solved puzzles in the non-erasure group was 4.50 ($SD = 1.35$, range = 0 to 6).

To test our main hypothesis that an erasure manipulation on the objective intelligence test data affects subjective intelligence ratings, a two-sided independent samples Bayesian t-test was conducted with experimental condition (erasure vs. non-erasure) as IV and subjective intelligence rating as DV. The Bayesian independent sample t-test (two-tailed, $N = 2{,}109$) yielded a final $BF_{10} = 0.38$, indicating anecdotal evidence in support of H0. The mean intelligence rating score was nearly identical in the erasure group ($M = 66.00$, $SD = 16.34$) and the non-erasure group ($M = 65.96$, $SD = 15.38$). Figure 1 visualizes the sequential analysis of BF (sequential BF curve) across all participants in the order of testing. The graph shows the change in BF over time as additional participants were tested, i.e., as more and more evidence were included in the analysis (see Figure 1).

Additional Analysis

In an additional pre-registered analysis, we also calculated the correlation (Bravais-Pearson; frequentist) between objective IQ test performance and subjective intelligence rating score in the non-erasure condition. The Bravais-Pearson correlation yielded $r(1014) = .19$, $p < .001$,

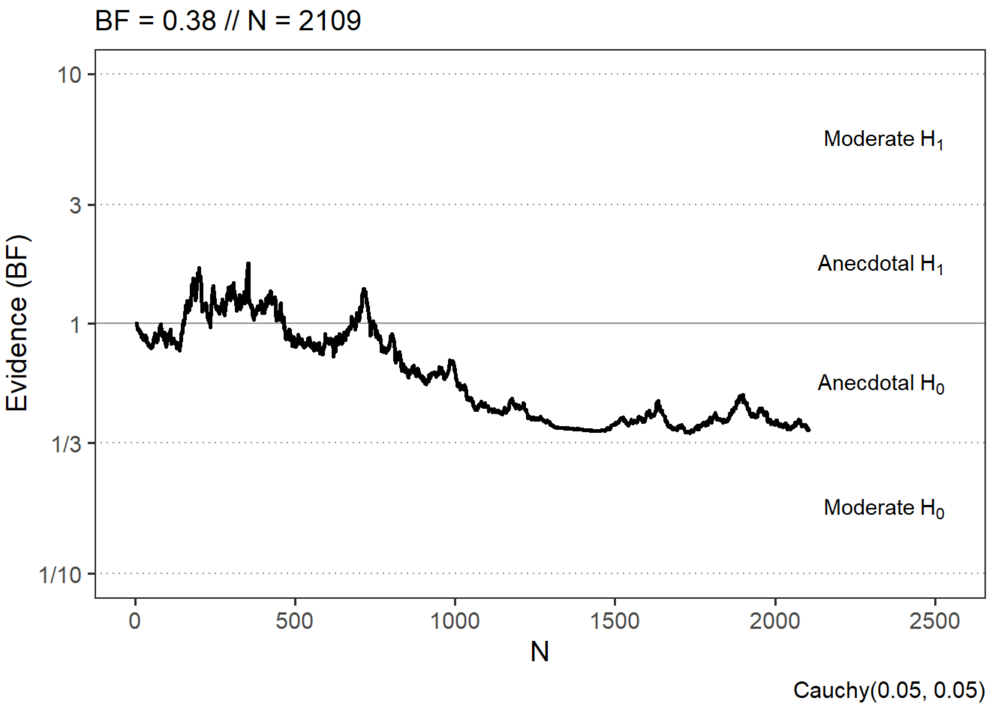

Figure 1. Sequential Bayes Factor Derived from the Independent Samples Bayesian T-Test Testing the Main Hypothesis Across Participants in Study 1.

indicating a predicted, albeit weak, positive relationship between subjective intelligence rating and objective IQ test performance.

Explorative Analyses

In addition to the pre-registered analyses reported above, a series of post-hoc analyses were conducted on the data to identify potential moderators of the erasure effect. Participants' gender was identified as one such potential moderator, as a general gender effect was observed on the subjective rating variable. Male participants rated themselves higher on subjective intelligence ($M = 68.97$, $SD = 16.97$) than female participants ($M = 64.35$, $SD = 14.97$), $t(2094) = 6.44$; $p < .001$. Based on this finding, we speculated that males may have an intentional bias toward overestimating and females an intentional bias toward underestimating their subjective intelligence. This would lead to an opposite trend when an erasure manipulation was administered. Within the male subsample, higher scores were expected in the erasure condition compared to the non-erasure condition, whereas within the female subsample, an opposite trend was predicted under the erasure manipulation.

To test this post-hoc assumption, a one-tailed Bayesian independent samples t-test with erasure manipulation (erasure vs. non-erasure) as the IV and subjective intelligence rating as the DV was first conducted within the male group, testing the hypothesis that males will have a higher mean subjective intelligence rating score in the erasure group compared to the mean score of males in the non-erasure group. This independent samples Bayesian t-test (one-tailed) yielded a final $BF_{10} = 1.35$, indicating anecdotal evidence for H1. The mean intelligence rating score was higher in the erasure group ($M = 69.66$, $SD = 17.79$) than in the non-erasure group ($M = 68.28$, $SD = 16.11$). The frequentist independent sample t-test revealed a weak trend in the predicted direction, $t(738) = 1.11$, $p = .13$. Next, a one-tailed Bayesian independent samples t-test was conducted within the female group with the erasure manipulation (erasure vs. non-erasure) as the IV and subjective intelligence rating as the DV, now testing the opposite hypothesis that females will have a lower mean subjective intelligence rating in the erasure group compared to the female mean in the non-erasure group. This independent samples Bayesian t-test (one-tailed) yielded a final $BF_{10} = 0.65$, indicating anecdotal evidence for H0. The mean intelligence rating score was lower in the erasure group ($M = 64.13$, $SD = 15.13$) than in the non-erasure group ($M = 64.59$, $SD = 14.79$). The frequentist analysis also showed no significant differences, $t = 0.57$. Although the final Bayesian evidence was inconclusive in both subsamples, gender could be considered a potential moderator of the erasure effects tested in this study.

Discussion

The focus of Study 1, reported above, was whether a subjective construction of reality, in this case operationalized by subjective ratings of one's intelligence on a visual analog scale, presumably constructed by volitional impulses, varied with the degree to which it was objectified. Theoretically, a more pronounced subjective bias toward volition was predicted in the erasure condition, in which no objectification of this reality through objective intelligence measures was available, and a smaller or no bias toward volition in the non-erasure condition, in which objectification through available objective intelligence data occurred simultaneously. In this way, the transferability of the complementarity principle to macroscopic psychological phenomena, as proposed by Bohr (1949a), was to be tested empirically, and a measurement-dependent description of subjective realities in the domain of intelligence assessment was to be investigated. Contrary to our pre-registered prediction, there was no statistically relevant difference in the means of subjective intelligence scores in the erasure condition compared to the non-erasure condition. The final Bayesian evidence does not allow a clear judgment on the validity of H1 or H0 in the available data. Thus, the intended evidence of a macroscopic complementarity relationship between subjective reality construction in the domain of intelligence assessment and its objective confirmation could not be provided. This result in itself calls into question the validity of the assumptions of Generalized Quantum Theory (Atmanspacher et al., 2002; Fach, 2011; Filk & Römer, 2011; Römer, 2023; Lucadou et al., 2007; Walach, 1998; Walach & Römer, 2000; Walach & Stillfried, 2011) and especially the Model of Pragmatic Information (Lucadou, 1984, 1987, 1995, 1998, 2001, 2002, 2015; Lucadou et al., 2007; Maier et al., 2022) for the psychological domain of volition.

However, the post-hoc gender differences found (see also Syzmanowicz & Furnham, 2011) might indicate a potential moderator variable. Participants' gender may have made an admittedly very weak contribution to the direction of erasure-dependent subjective-intentional reality construction in subjective intelligence ratings. Post-hoc analyses of the data in males and females considered separately provided ancedotal evidence that there was more subjective overestimation of intelligence in the male subsample in the erasure vs. non-erasure condition, whereas no effect or a slight trend in the opposite direction was found in females. Since these findings are in line with other research (Syzmanowicz & Furnham, 2011) they may indicate the existence of a moderator which should be further explored. Therefore, in the following Study 2, the hypothesized effect indicated post hoc for males was to be confirmed by means of a pre-registered study.

Study 2

In the following study, the objectification-dependent variation in men's subjective intelligence ratings found post hoc in the descriptive data was to be subjected to confirmatory empirical testing. The directed hypothesis this time was that men would report a higher mean subjective intelligence rating when the objective intelligence data were erased (erasure condition) than when the objective data were not erased (non-erasure condition). Confirmation of this hypothesis would support the principle of complementarity between subjective reality construction and objective reality mapping postulated in Generalized Quantum Theory for the field of volitional psychology. The study has been pre-registered online at OSF at https://osf.io/xty8a

Methods

Ethical Guidelines

Participants were informed of the voluntary nature of participation and of the data protection regulations (informed consent procedure). All data were stored in encrypted form and analyzed anonymously. The study was approved by the Ethics Committee of the Department of Psychology, LMU Munich.

Sample and Data Collection

Participants were recruited entirely online via the crowdsourcing platform www.prolific.com. People from Germany, Austria, Switzerland, and the United Kingdom participated in the study. We used Prolific's pre-sorting option to ensure that only male participants were included. Compensation for participation was GBP 1 per participant, which was also reimbursed through Prolific. A Bayesian sequential data analysis design was used. This approach allows for cumulative data collection and analysis, i.e. additional participants are tested and results are added to the dataset until a given Bayes factor for H1 (alternative hypothesis) or H0 (null hypothesis) is reached. In the present study, an a priori threshold of BF = 10 was set and pre-registered, i.e., data are collected until BF = 10 for H0 or H1 is reached. When this occurs, data collection is stopped. If a BF > 10 is not reached within the previously defined $N = 2{,}000$, data collection is also stopped (Bayesian sequential design with maximum N). The "stopping rule" BF = 10 for H0 or H1 was reached in the experimental data with a number of $N = 255$ participants. Data collection was online for a total of 2 days. As pre-registered, participants were excluded if they answered more than two items in less than five seconds ($n = 2$). Also excluded from the analysis were participants who indicated that they had not completed the test in a quiet, undisturbed environment ($n = 2$) or who indicated that their answers were not reliable ("use-me-item", $n = 1$). In addition, four individuals reported that they identified as female and were also excluded. The final sample included in the analysis consisted of $N = 246$ all-male participants from Germany, Austria, Switzerland, and the United Kingdom. The mean age of the participants was $M = 39.61$ years ($SD = 13.81$). Data collection could be completed online via PC, tablet, or smartphone, and participants could choose to complete the study in German or English.

Material, Experimental Procedure and Design

The present study was a direct replication of study 1 with the exception that only men were allowed as participants. Material, experimental procedure and design were therefore exactly identical to Study 1.

Results

Data collection and analysis were performed using Bayesian inference. We used an informed prior specified in the preregistration in Study 2 that follows a Cauchy distribution centered around 0.1, with $r = 0.05$, i. e., $\delta \sim$ Cauchy (0.1, 0.05). This distribution is based on a robustness analysis of the data from the male participants in the aforementioned Study 1.

The statistical hypothesis was that the mean subjective self-report of intelligence in the group in which the objective data were deleted would be higher than the mean subjective intelligence rating in the group of participants whose objective data were not deleted. A one-tailed Bayesian t-test for independent samples yielded a final $BF_{10} = 10.51$. The group mean in the erasure condition ($n = 117$) was $M = 67.56$ ($SD = 12.91$), and in the non-erasure condition ($n =$

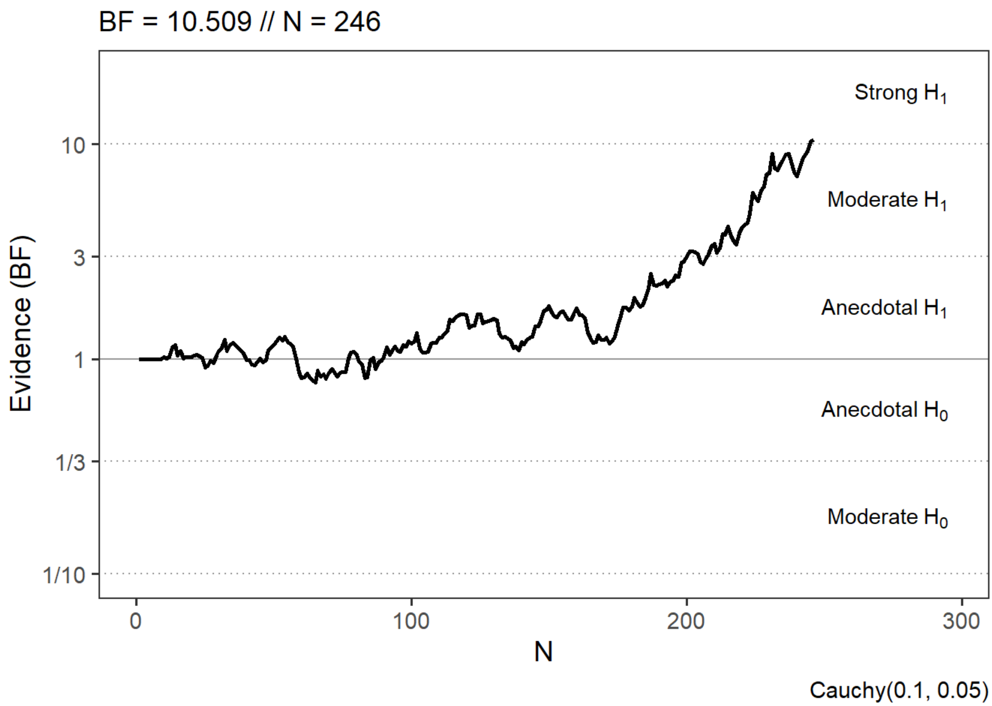

Figure 2. Sequential Bayes Factor of the Main Analysis in Study 2.

129), $M = 62.01$ ($SD = 17.79$). This can be interpreted as strong evidence for the validity of H1. Fig. 2 shows the sequential course of the Bayes factor.

Additional Analyses

The mean score of correctly solved items in the IQ test of the participants (objective data) whose results were stored (n = 146) was 3.64 out of 6 items. There was a range from no items correctly completed (minimum) to all six items correctly completed (maximum). The correlation between the objective intelligence data and the subjective scores in the non-erasure condition was $r(127) = .041$, $p = .64$., which was surprisingly low.

Discussion

Study 2 was designed to replicate and confirm the anecdotal post hoc erasure findings found in Study 1. It was expected that men's subjective intelligence ratings would be higher in the erasure condition than in the non-erasure condition. This was indeed demonstrated with strong evidence. This pattern of findings provides preliminary evidence for the complementary relationship between subjective reality construction and its objective confirmation in macroscopic data, and thus for the validity of one of the central assumptions of Generalized Quantum Theory (Atmanspacher et al., 2002; Fach, 2011; Filk & Römer, 2011; Römer, 2023; Lucadou et al., 2007; Walach, 1998; Walach & Römer, 2000; Walach & Stillfried, 2011) and especially the Model of Pragmatic Information (Lucadou, 1984, 1987, 1995, 1998, 2001, 2002, 2015; Lucadou et al., 2007; Maier et al., 2022). A necessary precondition for such evidence seems to be the normative direction of the intentional subjective bias, which was uniformly toward overestimation in the group of men. In the case of heterogeneous subjective bias, as in the case of a mixed-gender group (see Syzmanowicz & Furnham, 2011), such an effect would not be documented. However, a closer look reveals a surprising detail. The effect sizes appear to differ strongly between Study 1 and Study 2 (see Table 4). This effect heterogeneity could indicate additional unidentified moderators or future problems in replicating the effect. To address this issue, another pre-registered study (Study 3) was conducted to confirm the results found in Study 2.

Study 3

The present study should replicate the central finding of Study 2 to test the credibility of the effect documented there. This is somewhat doubtful because of the strong heterogeneity of the effect sizes of the erasure effects among men in Studies 1 and 2. The current study was also pre-registered at OSF (https://osf.io/fjhxy). Again, the research hypothesis was that men in the erasure condition would have higher mean subjective intelligence ratings than in the non-erasure condition.

Methods

Sample and Data Collection

Participants were again recruited online via the crowdsourcing platform www.prolific.com. Male participants were recruited from Germany, Austria, Switzerland and the UK. Compensation for participation was £1 for each participant and was also paid to them via Prolific. A Bayesian sequential data analysis design was again used. With the same stopping rules as in Study 2, a maximum N of 2,000 participants was set in case none of the stopping rules were reached; the actual N of usable data was $N = 1{,}986$. As pre-registered and analogous to Study 2, participants who answered more than two items in less than five seconds were excluded ($n = 17$). In addition, participants who reported that they did not take the test in a quiet, undisturbed environment ($n = 12$) or who reported that their responses were not reliable ("use-me-item", $n = 13$) were excluded from the analysis. In addition, 26 individuals reported that they identified as female and were also excluded. The final sample included in the analysis consisted of $N = 1{,}922$ all-male participants from Germany, Austria, Switzerland, and the United Kingdom. The mean age of the participants was $M = 38.90$ years ($SD = 13.16$). Data collection could be completed online via PC, tablet, or smartphone, and participants could choose to complete the study in German or English.

Material, Experimental Procedure and Design

The present study was a direct replication of Study 2, so the materials, experimental procedure, and design were exactly identical to Study 2.

Results

The data analysis was again Bayesian. An informed prior, specified in the preregistration, following a Cauchy distribution centered around 0.2 with $r = 0.15$, i. e., $\delta \sim$ Cauchy (0.2, 0.15) was used. This distribution is based on the effect size of $d_{Cohen} = .35$ reported in Study 2.

The statistical hypothesis was that the mean subjective self-report of intelligence in the group in which the objective data were deleted would be higher than the mean subjective intelligence rating in the group of participants whose objective data were not deleted. A one-tailed Bayesian t-test for independent samples yielded a final $BF_{01} = 1.90$. The mean of the group in the erasure condition ($n = 935$) was $M = 64.78$ ($SD = 15.25$), and in the non-erasure condition ($n = 987$) was $M = 64.27$ ($SD = 15.66$). This can be interpreted as anecdotal evidence for H0. Figure 3 shows the sequential course of the Bayes factor.

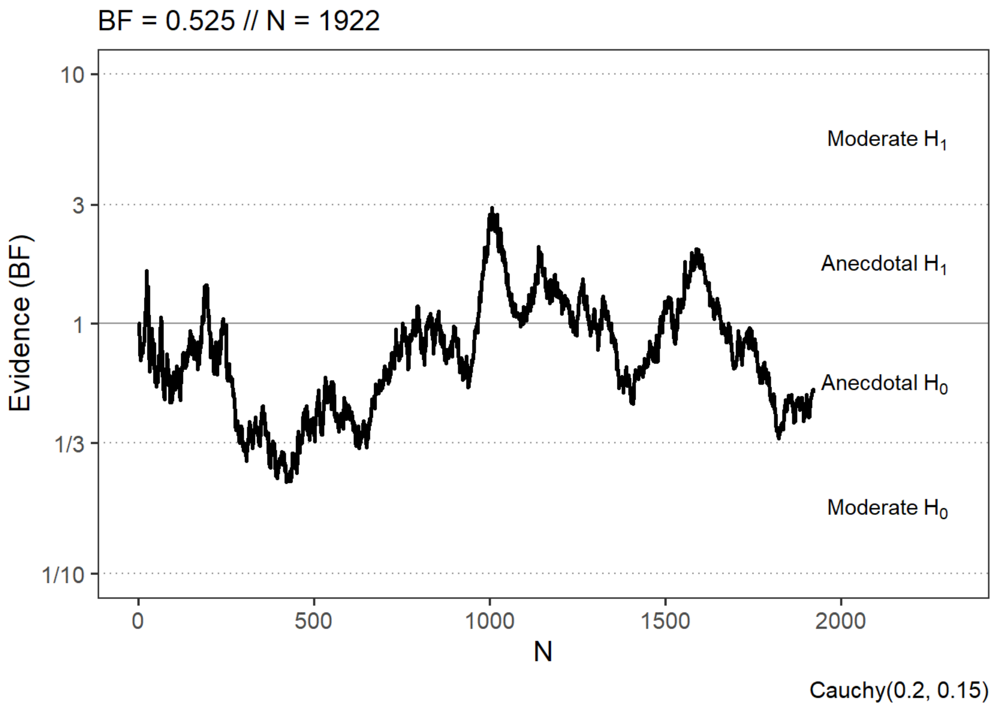

Figure 3. Sequential Bayes Factor of the Main Analysis in Study 3.

Additional Analyses

The mean number of correctly solved items in the IQ test of the individuals (objective data) whose results were stored (n = 987) was 3.70 out of 6 items. There was a range from no items correctly solved (minimum) to all six items correctly solved (maximum). The correlation between the objective intelligence data and the subjective rating in the non-erasure condition was $r(985) = .10$, $p < .01$.

Discussion

This study was designed to confirm the results of Study 2 by direct replication. After approximately reaching the maximum N-criterion of 2,000, only anecdotal Bayesian evidence for the null hypothesis could be found, contrary to the prediction. Thus, the replication attempt must be considered a failure. Men's subjective intelligence ratings did not differ as a function of the erasure manipulation. This calls into question the validity of the results of Study 2 and thus also the assumption of a macroscopic complementarity between subjective intelligence ratings and their objective confirmation – as postulated in Generalized Quantum Theory (Atmanspacher et al., 2002; Fach, 2011; Filk & Römer, 2011; Römer, 2023; Lucadou et al., 2007; Walach, 1998; Walach & Römer, 2000; Walach & Stillfried, 2011) and in the Model of Pragmatic Information (Lucadou, 1984, 1987, 1995, 1998, 2001, 2002, 2015; Lucadou et al., 2007; Maier et al., 2022).

The question was whether an as yet unidentified moderator could explain this unstable pattern of results across studies. Data exploration revealed a promising subpopulation consisting of German-speaking men (from Germany, Austria, and Switzerland) over the age of 30. To test this, a post hoc analysis was conducted with this subgroup (German-speaking males over 30: $n = 180$). This analysis revealed strong evidence for H1, $BF_{10} = 10.82$. As expected, the mean subjective intelligence score was higher in the erasure group, $M = 66.00$ ($SD = 13.20$), than in the non-erasure group, $M = 60.17$ ($SD = 14.61$). This result would confirm the findings of Study 2 and limit the occurrence of a directed erasure effect as formulated in the hypothesis to this subpopulation (German-speaking men over 30 years of age). In the following Study 4, the postulated erasure effect was to be tested confirmatory again for this subpopulation.

Study 4

The aim of this study was to confirm the post-hoc erasure effect found in Study 3 in German-speaking men over 30 years of age. For this purpose, another preregistration was formulated that predicted the directed erasure effect from Study 2, with the hypothesis that the mean subjective intelligence rating would be higher in the erasure group than in the non-erasure group in this subpopulation (https://osf.io/6hx98). Thus, a combination of gender (males), age (≥ 30 years), and nationality (German, Austrian, and German-speaking Switzerland) was assumed to be a relevant moderator. In addition, because sample restriction was assumed to be indicative of other underlying moderators, a series of items (5-point rating scale) were also collected at the end of the study to assess motivational importance and attitudes toward intelligence. The items were created following Andrew Elliot's motivational theory of achievement goals (Elliot & Thrash, 2001) and reformulated to address motivational goals related to intelligence. The following figure shows the 5 goal items used (see Fig. 4).

The first item captures the individual meaning of intelligence as a performance-approach goal (PAP), the second as a mastery-avoidance goal (MAV), the third as a mastery-approach goal (MAP), and the fourth as a performance-avoidance goal (PAV). The last item is used to capture the general meaning of an individual's intelligence (INT). The relationship between the items and the erasure effect should be examined.

Please read the following statements and indicate how strongly they apply to you.

	not at all		somewhat		very much
It is important to me to be more intelligent than others.	●	●	●	●	●
I worry that I am not acting as intelligently as I would like.	●	●	●	●	●
I want to process tasks as intelligently as possible.	●	●	●	●	●
I fear looking less intelligent than others.	●	●	●	●	●
Intelligence is an important issue for me.	●	●	●	●	●

Figure 4. Items For Recording Individual Goals In The Area Of Intelligence.

Sample and Data Collection

Participants were again recruited online via the crowdsourcing platform www.prolific.com. Male participants were recruited from the German-speaking countries (Germany, Austria and German-speaking Switzerland). Compensation for participation was £1 for each participant and was also paid via Prolific. A Bayesian sequential data analysis design was again used. With the same stopping rules as in Study 3, a maximum N of 1,000 participants was set if none of the stopping rules were reached. As pre-registered and analogous to Studies 2 and 3, participants who answered more than two items under five seconds were excluded ($n = 1$). In addition, participants who reported that they had not completed the test in a quiet, undisturbed environment ($n = 1$) or who reported that their responses were not reliable ("use-me-item", $n = 1$) were excluded from the analysis. In addition, one person identified as female and another person reported that they were younger than 30 years of age. Both were also excluded. The final sample included in the analysis consisted of $N = 192$ exclusively male over 30, German-speaking participants from Germany, Austria, and Switzerland. The mean age of the participants was $M = 38.36$ years ($SD = 7.97$). Data could be collected online using a PC, tablet, or smartphone.

Material, Experimental Procedure and Design

The present study was a direct replication of Studies 2 and 3, so the material, experimental procedure, and design were exactly identical to those studies.

Results

Data analysis was performed using Bayesian methods. An informed prior, specified in the pre-registration, following a Cauchy distribution centered around 0.2 with $r = 0.15$, i. e., $\delta \sim$ Cauchy (0.2, 0.15), was used. This distribution is based on the effect size of $d_{Cohen} = .42$ found in the post hoc analysis in Study 3.

The statistical hypothesis for the above subgroup was that the mean subjective self-report of intelligence in the group in which the objective data were deleted would be higher than the mean subjective self-report of intelligence in the participants whose objective data were not deleted. A one-tailed independent samples Bayesian t-test yielded a final $BF_{10} = 0.082$. The group mean in the erasure condition ($n = 97$) was $M = 62.87$ ($SD = 17.54$), and in the non-erasure condition ($n = 95$) was $M = 67.49$ ($SD = 12.26$). This can be interpreted as strong evidence for H0. Figure 5 shows the sequential course of the Bayes factor.

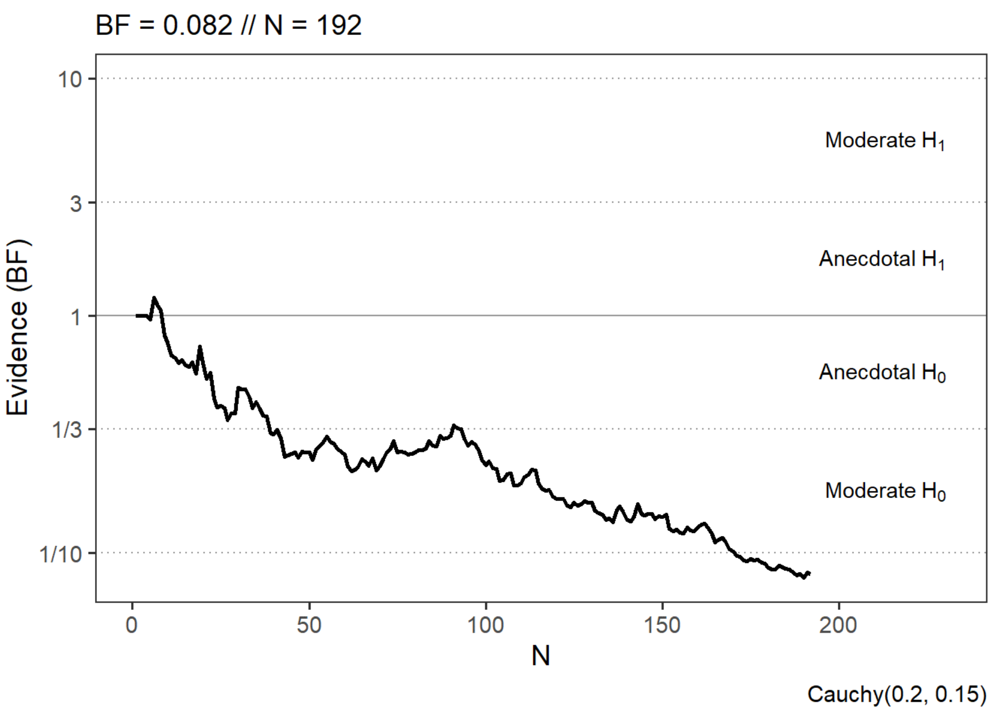

Figure 5. Sequential Bayes Factor of the Main Analysis in Study 4.

The mean of the objective intelligence data in the non-erasure condition was M = 4.47 (SD = 1.30), and the correlation (Bravais-Pearson) between subjective and objective intelligence was r = .15, p = .15.

Discussion

In the present study, we again attempted to replicate the erasure effect found in Study 2 among men in a subpopulation identified in Study 3 (German-speaking men over 30 years of age). Contrary to the pre-registered prediction, strong evidence for H0 emerged, again challenging the previous moderator assumptions and the theoretical models from which our predictions were derived. In fact, the effect tended to be the opposite. At this point in our research, we could either assume that the assumption of a macroscopic complementarity relationship was not valid for the area under investigation, or even in general, or that there were other, as yet unknown, moderators of the effect under investigation. The subjective intentional bias of a psychological concept such as self-reported intelligence could also vary within a homogeneous sample, assuming that participants recruited later have a more reserved opinion of their own intelligence and thus an underlying intentional tendency to underestimate it. This reluctance could correlate with the timing of their decision to participate, leading to a reversal effect at the end of a data collection period, as we observed in the present study. This admittedly very speculative assumption would imply an extreme context sensitivity of intentional biases, which is perfectly consistent with the subjective nature of the phenomenon under study.

In an additional post-hoc exploratory investigation, the influence of psychological factors (e. g., motivational importance of high intelligence, modesty in one's own intelligence assessment, etc.) was examined, as these might predict a direction of intentional bias and help to explain the partly contradictory sample-dependent effects. For the analysis, the participants in the study were divided into two groups characterized by a high and a low goal state for each goal item using a median split. The descriptive results of these post hoc analyses are presented in Table 1 below.

Subsample	Erasure		Non-Erasure	
	M (SD)	N	M (SD)	N
PAp high	70.9 (13.6)	53	69.6 (11.9)	67
PAp low	53.2 (17.0)	44	62.6 (11.8)	28
PAv high	62.0 (18.7)	80	66.6 (12.3)	79
PAv low	67.0 (9.57)	17	72.0 (11.4)	16
MAp high	64.9 (17.4)	68	68.8 (11.9)	77
MAp low	58.2 (17.2)	29	61.8 (12.6)	18
MAv high	66.2 (15.2)	69	69.2 (12.0)	74
MAv low	54.7 (20.4)	28	61.5 (11.7)	21
Int high	66.2 (15.2)	69	69.2 (12.0)	74
Int low	54.7 (20.4)	28	61.5 (11.7)	21

Note. P = Performance; M = Mastery; Ap = Approach; Av = Avoidance; Int = Importance of Intelligence.

Table 1. The Influence of Intelligence-Related Performance Goals and Meaningfulness of Intelligence (Median Split) on Erasure-Dependent Subjective Intelligence Assessment.

Study 5

The present study was conducted as an exploratory study to further investigate potential psychological moderator variables in the occurrence of an erasure effect on subjective intelligence ratings. The materials, experimental procedure, stimuli, and design were the same as in previous studies.

In addition, directed hypotheses were defined regarding the psychological variables: Individuals with a high approach goal (items 1 "PAP" and 3 "MAP") or a high stated importance of intelligence (item 5 "INT") should, on average, have a higher subjective intelligence rating in the erasure condition compared to the non-erasure condition, and individuals with low scores on these goals should have a reverse erasure effect.

Individuals with a high avoidance goal (items 2 "PAV" and 4 "PAP") should, on average, have a lower subjective intelligence rating in the erasure condition compared to the non-erasure condition, and individuals with low scores on these goals should have a reverse erasure effect.

In addition, anxiety was examined as another possible moderator. For this purpose, the short version of the State Scale of the Spielberger State-Trait Anxiety Inventory (STAIS-5; Zsido et al., 2020) was added to the questionnaire at the end of the study. The scale measures momentary anxiety with five items. Because the last of the 5 items ("I feel confused") seemed inappropriate because it might be too related to the previous matrix test, we replaced it with another item from the long version ("I feel worried"). The relationship between high and low anxiety and erasure-dependent subjective intelligence ratings was explored in a subsample (*N* = 95).

Methods

Sample and Data Collection

Participants were again recruited online via the crowdsourcing platform www.prolific.com. Participants (males and females) were recruited from the USA as previous pools had been exhausted. Compensation for participation was £1 per participant and was also paid via Prolific. A Bayesian sequential data analysis design was again used. The same stopping rules were used as in Study 3, except that no maximum N was set. Similar to Studies 2 and 3, participants who answered more than two tasks under five seconds were excluded (*n* = 11). Participants were also excluded from the analysis if they reported that they had not taken the test in a quiet, undisturbed environment (*n* = 4) or if they reported that their responses were not reliable ("use-me-item," *n* = 6). The final sample included in the analysis consisted of *N* = 375; 200 of the participants identified as male, 172 identified as female, and three identified as diverse. The mean age of the participants was *M* = 39.93 years (*SD* = 13.74). Data collection could be completed online using a PC, tablet, or smartphone.

Results

The directed hypotheses formulated above were tested using Bayesian t-tests for independent samples (informed prior $\delta \sim$ Cauchy (0.2, 0.15)), where the independent variable was the erasure manipulation (erasure vs. non-erasure condition) and the dependent variable was subjective intelligence rating using the visual analog scale. For each target group (PAPhigh, PAPlow, MAPhigh, MAPlow, PAVhigh, PAVlow, MAVhigh, MAVlow, and INThigh, INTlow), a one-tailed Bayesian t-test described above was performed separately, and depending on the target group, the direction of the predicted erasure effect varied as formulated in the hypotheses.

The results of the 10 analyses, along with the statistical hypotheses and descriptive statistics, are presented in the following table (see Table 2).

As can be seen from the table, hypothesized effects were only found for the PAP variable. Individuals with a PAPhigh target rated their subjective intelligence higher in the erasure condition than in the non-erasure condition, and a hypothesized opposite effect was found in the PAPlow group. No relevant Bayesian evidence for H1 was found for all other target groups.

Subsample	Hypothesis	Erasure (E)		Non-Erasure (NE)		BF_{10}
		M (SD)	N	M (SD)	N	
PAp high	E > NE	72.1 (13.5)	99	70.8 (14.1)	97	0.63
PAp low	E < NE	57.0 (18.2)	83	62.5 (17.3)	96	3.07
PAv high	E < NE	66.0 (16.1)	97	66.1 (14.7)	97	0.40
PAv low	E > NE	64.4 (18.9)	85	67.3 (17.7)	96	0.15
MAp high	E > NE	66.5 (17.0)	134	68.6 (15.2)	153	0.11
MAp low	E < NE	61.6 (18.3)	48	59.3 (18.1)	40	0.37
MAv high	E < NE	64.2 (15.6)	98	63.2 (17.0)	99	0.31
MAv low	E > NE	66.5 (19.4)	84	70.4 (14.6)	94	0.12
Int high	E > NE	67.5 (16.0)	143	67.3 (15.8)	153	0.27
Int low	E < NE	56.8 (20.0)	39	64.2 (17.8)	40	1.97

Note. P = Performance; M = Mastery; Ap = Approach; Av = Avoidance; Int = Importance of Intelligence.

Table 2. Results of Bayesian t-Tests for Independent Samples in Study 5.

Additional Analyses of Gender Differences in the Target Items

There were no significant gender differences in the target items.

Objective Intelligence and Correlation Between Subjective Assessment and Objective Intelligence

The mean of the objective intelligence data in the non-erasure condition was $M = 3.10$ ($SD = 1.70$) and the correlation (Bravais-Pearson) between subjective and objective intelligence was $r = 0$, $p = 1$.

State Anxiety as Moderator

Of the 95 participants whose State Anxiety was examined, only 11 exceeded the cut-off value of 9.5. In this group, there were descriptive differences between conditions (erasure: $M = 65.4$, $SD = 6.58$, $N = 5$ vs. non-erasure: $M = 55.7$, $SD = 12.3$, $N = 6$), but these were not statistically significant: $t(9) = 1.58$, $p = 0.15$. There were also no differences in the low anxiety group (erasure: $M = 64.9$, $SD = 23.5$, $N = 38$ vs. non-erasure: $M = 67.2$, $SD = 18.2$, $N = 46$; $t(82) = -0.52$, $p = .61$).

Discussion

The aim of the present study was to identify relevant psychological variables that predict the erasure effect in terms of an intentional bias toward overestimating or underestimating subjective intelligence. The PAP variable emerged as a promising candidate. Individuals with a high motivational tendency to be more intelligent than others (PAPhigh) seem to have an intentional bias toward a higher subjective estimate of their own intelligence, which varies accordingly with the use of the erasure manipulation (higher scores in the erasure condition vs. non-erasure condition), whereas individuals with a low expression of this variable (PAPlow) show an opposite trend and thus an intentional tendency toward underestimation, which again varies accordingly with the erasure manipulation. Thus, the PAP variable may provide the psychological explanation for the sample characteristic-dependent erasure effects found in the previous studies. Anxiety does not seem to be a variable that significantly moderates the effect.

Study 6

In the present study, we re-examined whether the psychological moderator PAP (high vs. low), which was found to be relevant in Study 5, could explain an opposite effect direction in subjective intelligence ratings as a function of the erasure manipulation. In addition, another psychological moderator variable was added in the current study to capture the intentional tendency to be modest in the context of one's own subjective intelligence assessment using an item. This item was added to the variables already used in Study 5 at the end of the study and reads, "I feel uncomfortable claiming to be very intelligent" (UNCOM). Anxiety was no longer included. The full list of items is shown in Figure 6. Methods, materials, and procedure were otherwise identical to Study 5.

Our hypothesis was that the PAP effects from Study 5 would be replicated. In addition, individuals who rated themselves as high on the UNCOM item (UNCOMhigh: ≥ 3) were expected to have lower mean subjective intelligence ratings in the erasure than in the non-erasure condition. In the UNCOMlow group, an opposite erasure effect was predicted. The explanation was: Individuals who are uncomfortable describing themselves as very intelligent have a deliberate tendency to report lower intelligence estimates, whereas individuals who are not uncomfortable should have a deliberate tendency to report higher subjective estimates.

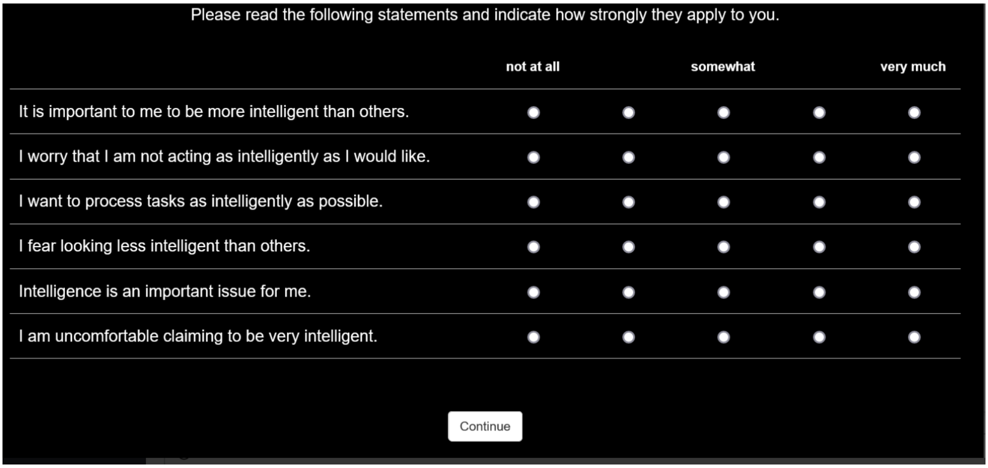

Figure 6. Items Assessing Individual Goals Regarding Intelligence and Discomfort in the Area of Intelligence Self-Assessment.

Methods

Sample and Data Collection

Participants were again recruited online via the crowdsourcing platform www.prolific.com. Participants (males and females) were recruited from the USA. Compensation for participation was £1 for each participant and was also paid to them via Prolific. A Bayesian sequential data analysis design was again used. The same stopping rules were used as in Study 3, again with no defined maximum N. A total of N = 588 participants were tested. Similar to the previous studies, participants who answered more than two items in less than five seconds were excluded (n = 12). In addition, participants who indicated that they had not taken the test in a quiet, undisturbed environment (n = 3) or who indicated that their responses were not reliable ("use-me-item," n = 5) were excluded from the analysis. The final sample included in the analysis consisted of N = 569; 316 of the participants identified as male, 243 identified as female, and 10 identified as diverse. The mean age of the participants was M = 42.46 years (SD = 13.79). Data collection could be completed online using a PC, tablet, or smartphone.

Results

The directed hypotheses formulated above were tested using Bayesian t-tests for independent samples (informed prior $\delta \sim$ Cauchy (0.2, 0.15)), where the independent variable was the erasure manipulation (erasure vs. non-erasure condition) and the dependent variable was the subjective intelligence rating using the visual analog scale. Specifically, this study focused on Bayesian analyses regarding the erasure effects for the PAP and UNCOM groups (high and low). For each target group (PAPhigh, PAPlow, MAPhigh, MAPlow, PAVhigh, PAVlow, MAVhigh, MAVlow, INThigh, INTlow, and UNCOMhigh, UNCOMlow) separately, a one-tailed Bayesian t-test as described above was performed, and depending on the target group, the direction of the predicted erasure effect varied as formulated in the hypotheses. The results of the 12 analyses, along with the statistical hypotheses and descriptive statistics, are presented in the following table.

Subsample	Hypothese	Erasure (E)		Non-Erasure (NE)		BF_{10}
		M (SD)	N	M (SD)	N	
PAp high	E > NE	72.1 (13.7)	159	73.2 (12.9)	146	0.13
PAp low	E < NE	57.2 (19.5)	143	59.7 (17.6)	121	0.83
PAv high	E < NE	65.2 (16.9)	157	67.1 (16.7)	133	0.73
PAv low	E > NE	64.9 (19.7)	145	67.0 (16.6)	134	0.12
MAp high	E > NE	67.4 (16.1)	227	69.3 (16.3)	208	0.08
MAp low	E < NE	57.7 (22.4)	75	59.1 (15.4)	59	0.55
MAv high	E < NE	63.5 (17.2)	158	65.2 (16.1)	136	0.66
MAv low	E > NE	66.7 (19.3)	144	69.0 (17.0)	131	0.11
Int high	E > NE	67.4 (16.5)	229	70.2 (14.7)	204	0.06
Int low	E < NE	57.5 (21.5)	73	57.0 (18.6)	63	0.40
Uncom high	E < NE	60.8 (16.9)	191	64.6 (16.1)	170	3.81
Uncom low	E > NE	72.4 (18.3)	111	71.3 (16.8)	97	0.48

Note. P = Performance; M = Mastery; Ap = Approach; Av = Avoidance; Int = Importance of Intelligence; Uncom = Uncomfortable Claiming to be Intelligent.

Table 3. Results Of Bayesian t-Tests for Independent Samples for the Individual Target Groups and the UNCOM Groups in Study 6.

As can be seen from the table, in contrast to the previous study, smaller or no hypothesis-consistent effects were found for the PAP variable. Individuals with a PAPhigh target even rated their subjective intelligence lower in the erasure condition than in the non-erasure condition, and a hypothesis-compliant inverse effect was shown only in the PAPlow group. In all other target groups, no relevant Bayesian evidence for H1 was found.

For the UNCOM groups, and especially for the UNCOMhigh group, a promising pattern of results emerged: moderate Bayesian evidence for H1 was found for UNCOMhigh participants. Figure 7 shows the sequential Bayesian progression, which consistently shows a positive trend toward H1.

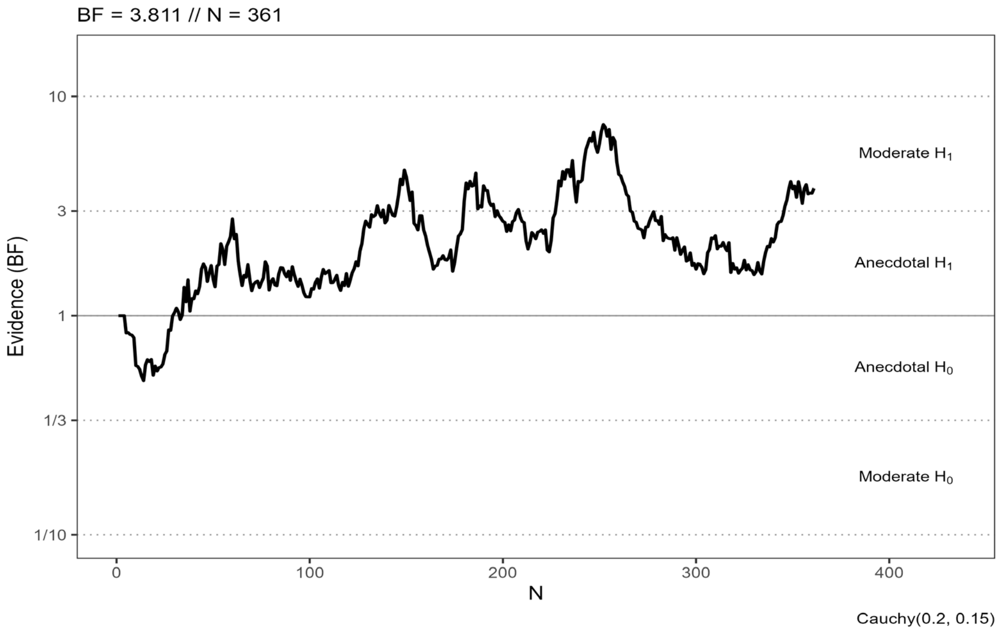

Figure 7. Sequential BF for the Erasure Effect (E < NE) in the UNCOMhigh Group in Study 6.

Additional Analyses of Sex Differences in the Target Items

There were significant gender differences in the following items:
PAP: $M_M = 2.72$ ($SD_M = 1.18$) and $M_F = 2.41$ ($SD_F = 1.13$); $t(557) = 3.18$, $p < .01$
MAV: $M_M = 2.49$ ($SD_M = 1.27$) and $M_F = 2.76$ ($SD_F = 1.24$); $t(557) = 3.18$, $p = .02$

Objective Intelligence and Correlation Between Subjective Assessment and Objective Intelligence

The mean of the objective intelligence data in the non-erasure condition was $M = 3.02$ ($SD = 1.66$) and the correlation (Bravais-Pearson) between subjective and objective intelligence was $r = .03$, $p = .58$.

Discussion

This study examined the role of PAP and UNCOM as potential psychological moderators of the direction of the erasure effect. For PAP, this was a replication of the results from Study 5, and for UNCOM, it was a new moderator that was examined for the first time. Although the results for the PAP variable did not meet expectations and therefore tended to lose weight as a possible moderator in the current results, a promising effect emerged for the UNCOM-high group. Individuals who feel uncomfortable describing themselves as very intelligent are likely to have a deliberate tendency to rate themselves lower in a required self-assessment. This tendency should be more pronounced when an erasure manipulation is used than when intelligence is objectified. The preliminary data support this hypothesis, although currently with only moderate evidence. In a future replication attempt (see Study 7), we will focus on this group (UNCOMhigh), as we believe it has the greatest potential to yield stable erasure effects and thus strong evidence for a macroscopic complementarity relationship between subjective intelligence assessment and objective confirmation.

Study 7

The aim of this study was to confirm the moderate erasure effect found in Study 6 within the UNCOMhigh subsample. For this purpose, the corresponding item "I feel uncomfortable claiming to be very intelligent" (UNCOM) was inserted on its own on a separate page after the assessment of subjective intelligence and used as a screening item. Only participants who reported a score of 3 ("somewhat") or higher were invited to participate in the remainder of the study. All other methods, materials, and procedures were otherwise identical to Study 5.

It was expected that participants who were uncomfortable describing themselves as intelligent would have a deliberate tendency to rate their subjective intelligence lower. This effect should come into play again in the erasure condition, where there is no objective test of intelligence. Thus, we hypothesize that among participants who complete the screening positively (UNCOM-high), the mean subjective intelligence estimate will be lower in the erasure condition than in the non-erasure condition. The study has been pre-registered at OSF (https://osf.io/k2ca7).

Methods

Sample and Data Collection

Participants were again recruited online via the crowdsourcing platform www.prolific.com. Participants (males and females) were recruited from the USA. Compensation for participation was £0.15 for participation in the screening and an additional £0.85 (i. e. a total of £1) for participation in the full study, and was paid through Prolific. A Bayesian sequential data analysis design was again used. Stopping rule was set to BF = 10 and maximum N = 2000 (sequential Bayesian analysis with maximum N). A total of $N = 1{,}670$ participants participated in the screening, of which $N = 1{,}042$ were assigned to the UNCOMhigh group and participated in the subsequent IQ test. Similar to previous studies, participants who answered more than two items in less than five seconds were excluded ($n = 26$). In addition, participants who indicated that they did not take the test in a quiet, undisturbed environment ($n = 12$) or who indicated that their answers were not reliable ("use-me-item," $n = 5$) were excluded from the analysis. The final sample included in the analysis consisted of $N = 1{,}002$; 440 of the participants identified as male, 542 identified as female, and 20 identified as diverse. The mean age of the participants was $M = 42.09$ years ($SD = 14.01$). Data collection could be completed online using a PC, tablet, or smartphone.

Results

The directed hypothesis formulated above was tested using a one-tailed Bayesian t-test for independent samples (informed prior δ ~ Cauchy (0.2, 0.15)), where the independent variable was the erasure manipulation (erasure vs. non-erasure condition) and the dependent variable was the subjective intelligence rating using the visual analog scale.

Evidence for the null hypothesis was found: subjective intelligence ratings were indeed higher in the erasure condition ($M = 63.61$, $SD = 16.95$) than in the non-erasure condition ($M = 62.86$, $SD = 16.58$). The BF_{10} was 0.15, reflecting moderate evidence for H0 (see Fig. 8). Although neither the threshold of BF ≥ 10 nor the maximum N were reached as stopping criteria, data collection was stopped at this point due to lack of funding. The current BF confirms with moderate evidence the null hypothesis with a Bayes factor of 6.80.

No sex differences were found.

Objective Intelligence and Correlation Between Subjective Assessment and Objective Intelligence

The mean of the objective intelligence data in the non-erasure condition was $M = 3.06$ ($SD = 1.61$) and the correlation (Bravais-Pearson) between subjective and objective intelligence was $r = .03$, $p = .46$.

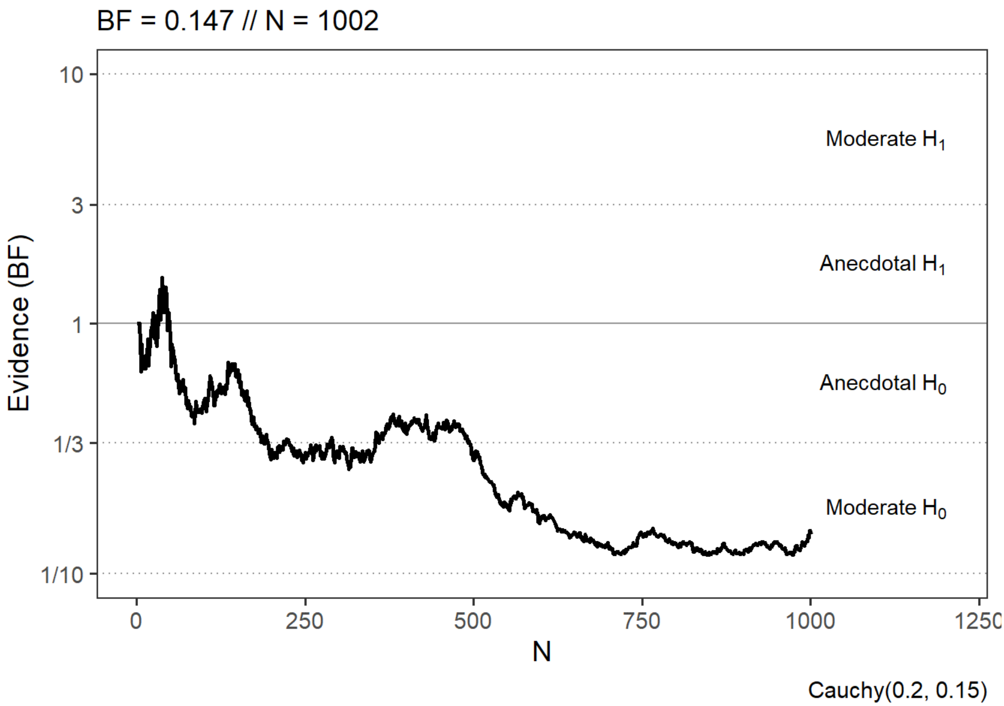

Figure 8. Sequential BF for the Erasure Effect (E < NE) in the UNCOMhigh Group in Study 7.

Discussion

Contrary to the hypothesis, no erasure effect on subjective intelligence ratings was found for the high UNCOM group. Although a strong intentional tendency to underestimate one's own intelligence was hypothesized for this group, and although post-hoc results in Study 6 supported this hypothesis, no difference between erasure groups was found in the present study. Despite all efforts to identify potential moderator variables that might underlie an intentional bias in subjective intelligence ratings, no replicable erasure effect, and thus no evidence of a macroscopic complementarity relationship, was found in the seven studies conducted. This raises the question of whether only subjective intelligence measures do not allow for stable, collective intentional tendencies, or whether this is a general phenomenon in the recording of subjective reality constructions. Further erasure studies with alternative reality constructions will investigate this question. Table 4 lists all results (BFs_{10} of the seven studies) together with the hypotheses and the effect sizes.

Study	(Sub-)Sample	Hypothesis	N	BF_{10}	p-Value	Cohens d
1	NA	E ≠ NE	2109	0.38	0.96	0.12
2	Male	E > NE	246	10.51	< .01**	0.35
3	Male	E > NE	1922	0.53	0.07	0.07
3	Male GAS	E > NE	180	19.82	< .01**	0.42
4	Male GAS >30 yrs	E > NE	192	0.08	0.98	-0.31
5	PAP high	E > NE	196	0.63	0.26	0.09
5	PAP low	E < NE	179	3.07	0.02*	0.31
5	PAV high	E < NE	194	0.40	0.48	0.01
5	PAV low	E > NE	181	0.15	0.86	-0.16
5	MAP high	E > NE	287	0.11	0.86	-0.13
5	MAP low	E < NE	88	0.37	0.72	-0.13
5	MAV high	E < NE	197	0.31	0.66	-0.06
5	MAV low	E > NE	178	0.12	0.93	-0.23
5	INT high	E > NE	296	0.27	0.46	0.01
5	INT low	E < NE	79	1.97	0.04*	0.39
6	PAP high	E > NE	305	0.13	0.76	-0.08
6	PAP low	E < NE	264	0.83	0.14	0.13
6	PAV high	E < NE	290	0.72	0.17	0.11
6	PAV low	E > NE	279	0.12	0.84	-0.12
6	MAP high	E > NE	435	0.08	0.89	-0.12
6	MAP low	E < NE	134	0.55	0.34	0.07
6	MAV high	E < NE	294	0.66	0.19	0.10
6	MAV low	E > NE	275	0.11	0.85	-0.13
6	INT high	E > NE	433	0.05	0.97	-0.18
6	INT low	E < NE	136	0.40	0.56	-0.02
6	UNCOM high	E < NE	361	3.81	0.01*	0.23
6	UNCOM low	E > NE	208	0.48	0.33	0.06
7	UNCOM high	E < NE	1002	0.15	0.76	-0.04

Note. E = erasure condition, NE = non-erasure condition; P = Performance; M = Mastery; Ap = Approach; Av = Avoidance; Int = Importance of Intelligence; Uncom = Uncomfortable Claiming to be Intelligent; GAS =German, Austrian or Swiss Participants.

Table 4. Overview of All Tests Performed. For a Detailed Presentation, see https://osf.io/c6kuw

General Discussion

The aim of the studies reported here was to find empirical evidence for a macroscopic complementarity relation between subjectively assessed, and thus intentionally biased, intelligence self-assessments and objective intelligence assessments of test participants. While Bohr originally proposed complementarity as applicable only to quantum phenomena, he speculated early on about its application to psychology (Bohr, 1929; 1938/1958) and recent theoretical developments have extended its scope to macroscopic domains, particularly in volitional psychology. The Generalized Quantum Theory (GQT) has provided a theoretical framework for this purpose, which at its core dispenses with the range of action of measurement-dependent phenomena restricted by Planck's quantum of action and thus extends such measurement-dependent complementarity relations to measurement arrangements describing macroscopic phenomena (Atmanspacher et al., 2002; Fach, 2011; Filk & Römer, 2011; Römer, 2023; Lucadou et al., 2007; Walach, 1998; Walach & Römer, 2000; Walach & Stillfried, 2011). In particular, psychophysical interactions between an individual's subjective experience or subjective-autonomous volition and objective-physical reality have been theoretically formalized in this way (Lucadou, 1984, 1987, 1995, 1998, 2001, 2002, 2015; Lucadou et al., 2007).

Despite these theoretical advancements, our empirical findings present mixed results. While one pre-registered study (Study 2) demonstrated an erasure-dependent effect between subjective intelligence ratings and objective assessments with strong Bayesian evidence, this finding could not be confirmed in replication attempts (Studies 3 and 4). Post-hoc analyses hinted at potential trends consistent with the theory, but their non-confirmatory nature limits their interpretative value.

Our interpretation of the data suggests that intentionally biased subjective assessments of one's own intelligence may be influenced by contextual and individual factors, leading to varying intentional biases in different directions (overestimation vs. underestimation). Therefore, identifying and understanding these factors becomes crucial for robustly replicating erasure effects in subjective intelligence assessment. Subsequent studies (5, 6, and 7) attempted to uncover such motivational goals underlying subjective intelligence ratings. While some promising leads emerged, particularly in Study 6 – where individuals who feel uncomfortable describing themselves as very intelligent showed a normative intentional tendency to bias their subjective intelligence ratings toward lower scores under erasure condition compared to non-erasure condition – these findings were not consistently replicated across studies.

The failure to replicate erasure effects raises questions about the existence of macroscopic complementarity relations in psychology. Lucadou's "no transmission" axiom suggests that attempting to objectify subjective phenomena ultimately leads to their destruction, since every subjective autonomous impulse turns into a normative-collective observable when confirmation attempts are made through replication. Objective confirmation thus ultimately destroys the subjective intentional impulse to create reality and leads to a random pattern of dispersion in the subjective variable. Without these limitations, the replicability of this type of phenomena would allow an experimenter to intentionally manipulate the measurement state of one classical observable (e. g., objective data), which would simultaneously bias in a normatively (non-randomly) controlled manner the entangled state of another classical observable (e. g., subjective data), providing instantaneous classical information transfer across time and space. However, Lucadou's argument does not invalidate GQT, but rather shows that any direct tests of macroscopic complementary relations by successful replications are strictly forbidden within the theoretical framework of GQT. This qualifies the GQT as a scientifically non-testable theory, at least within the scientific methodology currently accepted in the natural sciences. The main problem here is the impossibility of objectifying subjective phenomena with objective measurements while keeping the subjective nature of the concepts under study intact.

Let us apply this argument to our design: In attempting to manipulate intentionally biased subjective intelligence ratings with experimental erasure manipulations of their objective counterparts, demonstrating an erasure-dependent normative subjective rating effect would imply treating the subjective-autonomous domain exactly like a normative-objective bias. In other words, the objectification of subjective autonomous impulses would require the existence of a normative, quantitative group bias that contradicts or counteracts individual autonomy. Any objectification of macroscopic complementarities involving subjective effects will therefore lead to the destruction and disappearance of any subjective elements under study. In our view, this is the main reason for the failure to empirically document the objective existence of the macroscopic complementarity relations proposed in our studies. Similar replication failures in testing the replicability of macroscopic entanglement relations using the "matrix method" have recently been reported by others (Grote, 2021; Walach et al., 2022).

To resolve the paradox outlined above without abandoning the scientific method for testing macroscopic complementarity relations, new ways of testing these effects must be developed. One approach involves exploring the heterogeneity of intentional bias and testing it directly, rather than focusing solely on normative subjective bias. The heterogeneous subjective impulses reflected in individuals' self-reported intelligence scores should have more design latitude in a condition where objective data are absent than in a condition where objective data are present and the subjective-intentional moment is thereby constrained. This would suggest greater variability in subjective ratings in an erasure condition than in a non-erasure condition (we would like to thank Prof. Römer at this point for pointing out this possibility during a presentation discussion).

In addition, in another set of studies, instead of measuring the subjective autonomous impulse indirectly by assessing its effects on the concept under study with subjective ratings

(e. g., subjective intelligence ratings), a direct measure of the underlying process driving the intentional impulse could be implemented. Willpower, a concept widely used in motivational psychology, is one such candidate that could underlie the subjectively biased ratings and could be operationalized by direct measures, as suggested by Baumeister and Vohs (2007; ego-depletion) or Ryan and Deci (2008; vitality). In such studies, normative bias at the subjective level would not be a necessary condition for documenting macroscopic complementary relationships, but rather willpower would be the underlying unifying principle, independent of the individual rating, which could be autonomously biased in any direction (which would look like a random pattern). The directly assessed willpower underlying the subjective evaluation should then be higher in a condition where objective data are absent (erased or not assessed) than in a condition where objective data are present (measured and not erased).

In our future research on macroscopic complementarity relations, we will pursue both research approaches (variability tests and direct assessment of willpower) in addition to the previous investigation of normative erasure effects, in order to show that subjective, intentional reality construction and objective-deterministic reality descriptions are only two phenomenologically complementary, related versions of reality within a holistic worldview.

Finally, apart from the GQT related interpretations of our results, another potential explanation of the effects and their non-replicability across studies should be mentioned here. Our experimental tests of the complementarity relation between a subjectively created reality and variations of its objectification was based on subsample mean score measurements (subjective and objective summary assessments of individual scores). The complementarity principle proposed was thus not located at the level of the individual participant but on the level of the data analyst or experimenter of the studies. This would render experimenter psi (e-psi) a possible candidate for explaining our effects. According to this conjecture, the data could have been affected by such psi-like experimenter effects. Specifically, the assignment to the experimental condition (erasure vs. save condition) was performed by a pseudo-random process which could have unconsciously been affected by the investigators of the effect under study. This would reflect a mind-matter effect also known as micro-psychokinesis (micro-PK; for an overview see Varvoglis & Bancel, 2015). E-psi is a well know phenomenon in micro-PK research (e. g., Kennedy & Taddonio, 1976) and the PMIR model (Stanford, 1974) suggests that psi effects can arise unconsciously provided they fulfill an experimenter's need. Some studies have reported anecdotal and indirect evidence of e-psi in micro-PK (see Palmer, 2017) and the use of "silent" or "hidden" RNGs, which were concealed from the principal investigator during their studies, revealed significant results indicating e-psi (Berger, 1988; Honorton & Tremmel, 1979; Varvoglis, 1989; Varvoglis & McCarthy, 1986). It seems therefore possible that the effects reported in our studies and their variations reflected the unconscious beliefs or anxious disbeliefs of the investigators during the course of investigations. If this was the case, more effort needs to be spent in future studies to minimize potential e-psi effects, for example, by blinding conditions and/or by the involvement of uninformed data analysts. The latter might be especially relevant for Bayesian analyses were a continuous data monitoring was performed.

In addition, a further limitation, that affects most studies reported here, is the use of multiple Bayesian analysis for hypotheses testing within studies without correcting for multiple testing. In Bayesian statistics the Bayes Factor, our central testing score, needs not to be corrected when multiple testing occurs and can thus be interpreted with regard to evidence for H1 or H0 in the well-known manner without any corrections (see Jakob et al., 2024). However, as some suggest (e. g., de Jong, 2019; Westfall et al., 1997), the a priori prior should be adjusted depending on the number of tests performed ensuring a fair distribution of possible H1 and H0 interpretations. This would imply that our a priori model belief about the existence of macroscopic complementarity relations should be shifted more toward the skeptical end of our individual amount of subjective belief in the effect. Evidence for H1 would then push this after knowing the data into a less skeptical belief state in case evidence for H1 would have been found. Taken together, the shift in amount of belief in the effect reflected by the Bayes factor remains the same, only the absolute starting and end points concerning effect-belief would be changed by multiple testing control. Since, most of our effects document null findings and no real shift in our belief toward an existence of the effect could be found in the studies reported here, a multiple control analysis providing a correction of the a-priori prior was not reported here.

In conclusion, our research highlights the complexities of investigating macroscopic complementarity relations and underscores the importance of integrating theoretical insights with empirical evidence. While our findings present challenges, they also open new avenues for inquiry and provoke deeper reflection on the nature of subjective reality construction.

References

Alicke, M. D., & Sedikides, C. (2009). Self-enhancement and self-protection: What they are and what they do. *European Review of Social Psychology, 20,* 1–48. https://doi.org/10.1080/10463280802613866

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

Baumeister, R. F., & Vohs, K. D. (2007). Self-regulation, ego depletion, and motivation. *Social and Personality Psychology Compass, 1*(1), 115–128. https://doi.org/10.1111/j.1751-9004.2007.00001.x

Berger, R. E. (1988). Psi effects without real-time feedback. *Journal of Parapsychology, 52*(1), 1–27.

Bohr, N. (1928). The Quantum Postulate and the Recent Development of Atomic Theory. *Nature, 121*, 580–590. https://doi.org/10.1038/121580a0

Bohr, N. (1929). Wirkungsquantum und Naturbeschreibung. *Die Naturwissenschaften, 17*, 483–486. https://doi.org/10.1007/BF01505680

Bohr, N. (1948). On the notion of causality and complementarity. *Dialectica, 2*, 312–319. https://doi.org/10.1111/j.1746-8361.1948.tb00703.x

Bohr, N. (1949). *Discussions with Einstein on epistemological problems in atomic physics*. Cambridge University Press.

Bohr, N. (1985). Biologie und Atomphysik. In N. Bohr, *Atomphysik und menschliche Erkenntnis: Aufsätze und Vorträge aus den Jahren 1930 bis 1961* (pp. 12–21). Vieweg. https://doi.org/10.1007/978-3-322-88801-3

Brandstätter, V., Schüler, J., Puca, R. M., & Lozo, L. (2018). *Motivation und Emotion*. Springer. https://doi.org/10.1007/978-3-662-56685-5

de Jong, T. (2019). A Bayesian approach to the correction for multiplicity. https://doi.org/10.31234/osf.io/s56mk

Elliot, A. J., & Thrash, T. M. (2001). Achievement goals and the hierarchical model of achievement motivation. *Educational Psychology Review, 13*, 139–156. https://doi.org/10.1023/A:1009057102306

Engeler, I., & Häubl, G. (2021). Miscalibration in predicting one's performance: Disentangling misplacement and misestimation. *Journal of Personality and Social Psychology, 120*(4), 940–955. https://doi.org/10.1037/pspi0000301

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

Fahrenberg, J. (2013). *Zur Kategorienlehre der Psychologie: Komplementaritätsprinzip, Perspektiven und Perspektiven-Wechsel.* [On Category Systems in Psychology. Complementarity Principle, Perspectivism and Perspective-taking]. Pabst Science Publishers. https://doi.org/10.23668/psycharchives.10413

Favrholdt, D. (1999). *Complementarity beyond physics (1928–1962). Volume 10*. Elsevier.

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

Furnham, A. (2001). Self-estimates of intelligence: Culture and gender difference in self and other estimates of both general (g) and multiple intelligences. *Personality and Individual Differences, 31*(8), 1381–1405. https://doi.org/10.1016/S0191-8869(00)00232-4

Gignac, G. E., & Zajenkowski, M. (2019). People tend to overestimate their romantic partner's intelligence even more than their own. *Intelligence, 73*, 41–51. https://doi.org/10.1016/j.intell.2019.01.004

Grote, H. (2021). Mind–Matter entanglement correlations: Blind analysis of a new correlation matrix experiment. *Journal of Scientific Exploration, 35*, 287–310.

Heydasch, T., Haubrich, J., & Renner, K.-H. (2013). Die Kurzform des Hagener Matrizen-Tests (HMT-S): Ein 6-Item Intelligenztest zum schlussfolgernden Denken. *Methoden, Daten, Analysen, 7*, 183–208.

Heydasch, T., Haubrich, J., & Renner, K.-H. (2017). The short version of the Hagen Matrices Test (HMT-S): 6-Item Induction Intelligence Test. *Methods*, *Data, Analyses, 7*, 183–208. https://doi.org/10.12758/MDA.2013.011

Honorton, C., & Tremmel, L. (1979). Psi correlates of volition: A preliminary test of Eccles' "neurophysiological hypothesis" of mind-brain interaction. In W. G. Roll (Ed.), *Research in parapsychology 1978* (pp. 36–38). Scarecrow Press.

Jakob, J.-M., Dechamps, M. C., & Maier, M. A. (2024). Response to comment on Jakob, Dechamps, & Maier: How to read a paper. *Journal of Anomalous Experience and Cognition,* im Druck.

Kennedy, J. E., & Taddonio, J. L. (1976). Experimenter effects in parapsychological research. *Journal of Parapsychology, 40*(1), 1–33.

Lucadou, W. v. (1984). What is wrong with the definition of psi? *European Journal of Parapsychology, 5*, 261–283.

Lucadou, W. v. (1987). A multivariate PK experiment. Part III. Is PK a real force? The results and their interpretation. *European Journal of Parapsychology, 6,* 369–428.

Lucadou, W. v. (1995). The model of pragmatic information (MPI). *European Journal of Parapsychology, 11*, 58–75.

Lucadou, W. v. (1998). The exo-endo-perspective of non-locality in psycho-physical systems. *International Journal of Computing Anticipatory Systems, 2,* 169–185.

Lucadou, W. v. (2001). Hans in luck: The currency of evidence in parapsychology. *Journal of Parapsychology, 65,* 3–16.

Lucadou, W. v. (2002). Die Magie der Pseudomaschine. In W. Belschner, J. Galuska, H. Walach & E. Zundel (Eds.), *Transpersonale Forschung im Kontext* (pp. 77–100). Oldenburg.

Lucadou, W. v. (2015). The correlation-matrix method (CMM) – A new light upon the repeatability problem of parapsychology. *Journal of Parapsychology, 49*, 145–146.

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

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

Neubauer, A. C., & Hofer, G. (2020). Self- and other-estimates of intelligence. In R. J. Sternberg (Ed.), *Cambridge Handbook of Intelligence* (2nd ed., pp. 1179–1200). Cambridge University Press. https://doi.org/10.1017/9781108770422.050

Palmer, J. (2017). 'Experimenter Effects'. *Psi Encyclopedia*. London: The Society for Psychical Research. <https://psi-encyclopedia.spr.ac.uk/articles/experimenter-effects>. Retrieved 10 June 2021.

Römer, H. (2023). Twenty years of generalized quantum theory. *Journal of Anomalistics, 23*, 145–154.

Rudolph, U. (2013). *Motivationspsychologie kompakt*. Beltz.

Ryan, R. M., & Deci, E. L. (2008). From ego depletion to vitality: Theory and findings concerning the facilitation of energy available to the self. *Social and Personality Psychology Compass, 2*(2), 702–717. https://doi.org/10.1111/j.1751-9004.2008.00098.x

Stanford, R. G. (1974). An experimentally testable model for spontaneous psi events: I. Extrasensory events. *Journal of the American Society for Psychical Research, 68*, 34–57.

Syzmanowicz, A., & Furnham, A. (2011). Gender differences in self-estimates of general, mathematical, spatial and verbal intelligence: Four meta analyses. *Learning and Individual Differences, 21*, 493–504. https://doi.org/10.1016/j.lindif.2011.07.001

Varvoglis, M. (1989). A "psychic contest" using a computer-RNG task in a non-laboratory setting. In L. A. Henkel & R. E. Berger (Eds.), *Research in parapsychology 1997*. Scarecrow.

Varvoglis, M., & Bancel, P. A. (2015). Micro-psychokinesis. In E. Cardeña, J. Palmer & D. Marcusson-Clavertz (Eds.), *Parapsychology: A handbook for the 21st century.* (pp. 266–281). McFarland & Co.

Varvoglis, M. P., & McCarthy, D. (1986). Conscious-purposive focus and PK: RNG activity in relation to awareness, task-orientation, and feedback. *Journal of the American Society for Psychical Research, 80*(1), 1–29.

Walach, H., Kirmse, K. A., Sedlmeier, P., Vogt, H., Hinterberger, T., & Lucadou, W. v. (2022). Nailing jelly: The replication problem seems to be unsurmountable: Two failed replications of the matrix experiment. *Journal of Scientific Exploration*, *35*(4), 788–828. https://doi.org/10.31275/20212031

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

Walach, H., & Stillfried, N. v. (2011). Generalised 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

Weizsäcker, E. v. (1974). Erstmaligkeit und Bestätigung als Komponenten der pragmatischen Information. In E. v. Weizsäcker (Ed.), *Offene Systeme I* (pp. 82–113). Klett-Cotta.

Westfall, P. H., Johnson, W. O., & Utts, J. M. (1997). A Bayesian perspective on the Bonferroni adjustment. *Biometrika, 84*(2), 419–427. https://doi.org/10.1093/biomet/84.2.419

Zeh, H.-D. (1970). On the interpretation of measurement in quantum theory. *Foundations of Physics*, *1*, 69–76. Reprinted in Wheeler & Zurek (1983), pp. 342–349.

Zsido, A. N., Teleki, S. A., Csokasi, K., Rozsa, S., & Bandi, S. A. (2020). Development of the short version of the Spielberger State—Trait Anxiety Inventory. *Psychiatry Research*, *291*, 113223. https://doi.org/10.1016/j.psychres.2020.113223