Vaternam® Research

RESEARCH PAPER

Das Mittelmaß-Paradox

Generative KI hebt die Qualität – und macht strategische Markenentscheidungen wertvoller

Wenn Ausführung zur Commodity wird, wandert der Wettbewerbsvorteil vom Erzeugen zum Entscheiden.

HERAUSGEBER	Vaternam® Research · Vaternam® Kreativagentur e.K.
STAND	24. Juli 2026 · Version 1.0
FOKUS	Generative KI · Markenstrategie · kreative Vielfalt · Governance

VATERNAM RESEARCH PAPER

Inhalt

Dokumenttyp: strukturierte Evidenzsynthese, keine PRISMA-systematische Übersichtsarbeit. Stichtag der Recherche: 24. Juli 2026.

Abstract

Die verbreitete Behauptung, generative Künstliche Intelligenz produziere vor allem schlechte Inhalte, hält einer Prüfung der Forschung nicht stand. Kontrollierte Experimente und Feldstudien zeigen vielmehr, dass generative KI die Bearbeitungszeit vieler Schreib-, Beratungs-, Service- und Gestaltungsaufgaben verkürzt und die durchschnittlich bewertete Qualität häufig erhöht. Zugleich dokumentiert eine wachsende Forschungslinie ein anderes Risiko: Einzelne Ergebnisse werden besser, während Ideenpools, Geschichten, Essays und Modellantworten einander ähnlicher werden. Das relevante Problem ist daher nicht notwendig Qualitätsverfall, sondern eine Kombination aus drastisch wachsender Produktionsmenge, angehobenem handwerklichem Mindestniveau und potenziell sinkender kollektiver Varianz.

Dieses Paper entwickelt aus 39 ausgewählten wissenschaftlichen und institutionellen Quellen eine strukturierte Evidenzsynthese für die Markenführung. Es verbindet Forschung zu KI-Produktivität und kreativer Homogenisierung mit etablierten Theorien zu Markenwissen, glaubwürdigen Signalen, strategischen Trade-offs, Werbewirkung und Informationsüberlastung. Die Synthese ergibt eine präzisere These: Wenn die Grenzkosten akzeptabler Kommunikation fallen, verlagert sich der Engpass von der Erzeugung zur strategischen Auswahl. Wertvoller werden insbesondere Entscheidungen darüber, welche Bedeutung eine Marke besetzt, welche Belege sie autorisieren kann, welche Ausdrucksformen sie wiederholt und welche Themen, Versprechen, Zielgruppen oder Verhaltensweisen sie bewusst ausschließt.

Das Paper bezeichnet diesen abgeleiteten Effekt als **strategische Selektionsprämie**. Gemeint ist kein bereits kausal gemessener Geldwert, sondern ein theoretisch begründeter Hebel: Je mehr Ausgaben eine Entscheidung steuert, desto größer werden ihr Nutzen bei richtiger Wahl und ihr Schaden bei falscher Wahl. Das vorgeschlagene Vaternam-Modell trennt deshalb menschlich verantwortete Markenkonstitution, KI-gestützte Exploration und Produktion sowie ein evidenzbasiertes redaktionelles Gate. Die Schlussfolgerung lautet nicht „Mensch gut, KI schlecht“, sondern: KI ist besonders wertvoll innerhalb klarer strategischer Grenzen; sie ist besonders riskant, wenn sie diese Grenzen aus allgemeinen Mustern selbst ableiten soll.

Schlüsselwörter: generative KI; Markenstrategie; Positionierung; kreative Vielfalt; Homogenisierung; Markenführung; Distinktheit; Authentizität; Mensch–KI-Zusammenarbeit

Zentrale Aussage

> **Das Problem generativer KI ist nicht, dass sie nur schlechte Inhalte produziert. Das Problem ist, dass sie akzeptable und teilweise sehr gute Inhalte in einer Menge produziert, in der Güte allein kaum noch unterscheidet.**
>
> **Wenn Ausführung zur Commodity wird, wandert der Wettbewerbsvorteil vom Erzeugen zum Entscheiden.**
>
> **Eine Marke ist nicht die Summe dessen, was sie sagen kann. Sie ist die Konsistenz dessen, was sie auswählt, belegt und glaubwürdig ausschließt.**

Diese drei Sätze sind keine Einzelbefunde einer Studie. Sie verdichten die in diesem Paper entwickelte Synthese. Ihre wissenschaftliche Tragfähigkeit hängt von einer sauberen Unterscheidung ab: zwischen durchschnittlicher Ausführungsqualität, individueller Neuartigkeit, kollektiver Vielfalt, strategischer Eigenständigkeit und Markenwirkung.

1. Die falsche Frage

Die öffentliche Debatte über KI-Inhalte kreist häufig um die Frage, ob Maschinen „kreativ“ seien oder nur Mittelmaß erzeugten. Für Marken ist diese Frage zu grob. Sie vermischt mindestens vier verschiedene Größen: die handwerkliche Qualität eines einzelnen Ergebnisses, dessen Neuartigkeit, die Vielfalt einer Menge von Ergebnissen und die strategische Passung zu einer konkreten Organisation. Ein Text kann grammatisch, plausibel und angenehm lesbar sein, ohne eine unterscheidbare Markenentscheidung zu verkörpern. Ein Motiv kann hohe Klickwerte erzielen, ohne langfristig zum Aufbau eindeutiger Markenassoziationen beizutragen. Umgekehrt kann ein ungewöhnlicher Ausdruck neuartig, aber irrelevant oder unglaubwürdig sein.

Die Formulierung „KI produziert unendlich viel Mittelmaß“ ist deshalb als publizistische Zuspitzung brauchbar, wissenschaftlich aber in zwei Punkten zu korrigieren. Erstens ist die Produktion nicht unendlich, sondern praktisch sehr stark skalierbar: Sie bleibt durch Rechenleistung, Zugang, Energie, Zeit, Rechte und organisatorische Prozesse begrenzt. Zweitens ist der Output nicht notwendig mittelmäßig im Sinne geringer Qualität. Zahlreiche Studien zeigen Qualitätsgewinne. Das hier untersuchte Mittelmaß ist präziser als **strategische Austauschbarkeit** zu definieren: Ein Ergebnis erfüllt übliche Qualitätsanforderungen, könnte aber mit geringen Änderungen ebenso von einem Wettbewerber stammen, weil es keine proprietäre Wahrheit, keinen verbindlichen Trade-off und keine wiedererkennbare Auswahlentscheidung enthält.

Damit verschiebt sich die Forschungsfrage. Nicht: „Kann KI gute Kommunikation erzeugen?“ Das kann sie in vielen klar umrissenen Aufgaben. Sondern: **Was geschieht mit dem Wert strategischer Markenentscheidungen, wenn akzeptable Kommunikation schnell, billig und massenhaft verfügbar wird?**

Die Antwort dieses Papers lautet: Der Wert dieser Entscheidungen sinkt nicht automatisch mit den Produktionskosten. Unter plausiblen Bedingungen steigt ihr Hebel. Diese Aussage ist eine interdisziplinäre Schlussfolgerung, kein direkt beobachteter Kausaleffekt. Keine der identifizierten Studien schätzt den monetären Grenzwert einer Markenpositionierung vor und nach Einführung generativer KI. Die These wird vielmehr aus vier robusten Befundgruppen abgeleitet:

1. Generative KI reduziert Zeit und Kosten vieler Kommunikationsaufgaben und erhöht häufig die durchschnittliche Qualität.
2. KI-Unterstützung kann Leistungsunterschiede zwischen Personen komprimieren und Ergebnisverteilungen konvergenter machen.
3. Bei kreativen Aufgaben findet sich wiederholt ein Auseinanderfallen von besserer Einzelbewertung und geringerer kollektiver Vielfalt.
4. Markenwert beruht nach etablierter Forschung nicht allein auf guter Ausführung, sondern auf starken, günstigen und einzigartigen Assoziationen, glaubwürdigen Signalen, Wiedererkennbarkeit und strategischer Konsistenz.

2. Methode und Evidenzgrenzen

2.1 Design der Untersuchung

Das Paper ist eine **strukturierte Evidenzsynthese**, keine systematische Übersichtsarbeit im PRISMA-Sinn und keine Metaanalyse. Recherchiert wurden bis zum Stichtag 24. Juli 2026 peer-reviewte Artikel, Metaanalysen, randomisierte und quasi-experimentelle Studien, Feldstudien, etablierte theoretische Arbeiten, offizielle Forschungsberichte und – klar gekennzeichnet – einzelne Working Papers beziehungsweise Preprints.

Die Suche verband Begriffe aus fünf Literaturfeldern: generative AI und Produktivität; Human–AI Collaboration und Urteil; Kreativität, Neuartigkeit, Vielfalt und Homogenisierung; Markenstrategie, Markenwissen, Glaubwürdigkeit und Distinktheit; Informationsüberlastung und Aufmerksamkeit. Verwendet wurden Fachverlagsseiten, DOI-Register und öffentlich zugängliche wissenschaftliche Repositorien, darunter Science, Nature, Oxford Academic, INFORMS, Sage, Springer, Elsevier/RePEc, ACM, PubMed/PMC, arXiv, SSRN, OSF sowie institutionelle Publikationsseiten von Stanford, MIT, Harvard, Microsoft Research und OECD. Ausgangspublikationen wurden durch Rückwärts- und Vorwärtssuche ergänzt.

Die Auswahl priorisierte:

- direkte empirische Relevanz zur Forschungsfrage;
- transparente Methoden und identifizierbare Stichproben;
- peer-reviewte und replizierbare Befunde;
- Gegenbefunde, die einer einfachen „KI macht alles schlechter“-Erzählung widersprechen;
- internationale oder feldnahe Evidenz, soweit verfügbar;
- grundlegende Marken- und Strategietheorien, deren Konstrukte den Transfer zur Markenführung tragen.

Popularisierende Marktberichte wurden nicht als Belege für Wirkungszusammenhänge verwendet. Aktuelle Adoptions- und Kostendaten stammen aus dem Stanford AI Index und werden als Berichts- beziehungsweise Umfragedaten behandelt, nicht als experimentelle Evidenz.

2.2 Was „weltweit“ hier bedeutet

Eine Einbeziehung „aller verfügbaren Studien weltweit“ ist methodisch nicht seriös zusicherbar. Publikationen erscheinen fortlaufend, Datenbanken unterscheiden sich in Abdeckung und Sprache, und

ein Teil der Forschung ist nicht öffentlich indexiert. Zudem ist die aktuelle GenAI-Literatur überproportional englischsprachig und auf digitale Wissensarbeit in Nordamerika, Europa und Teilen Asiens konzentriert.

Der Anspruch dieses Papers ist daher **breit, aktuell, überprüfbar und widerspruchssensibel**, nicht buchstäblich vollständig. Die internationale Übertragbarkeit wird dort gestützt, wo Studien mehrere Länder oder globale Arbeitskontexte umfassen. Sie bleibt dort eingeschränkt, wo Befunde aus einzelnen Plattformen, Unternehmen, Aufgaben oder Kulturen stammen.

2.3 Evidenzgrade

Zur Einordnung werden vier Evidenzgrade verwendet:

- **A:** peer-reviewte Metaanalyse, systematisches Review oder große Feldstudie mit belastbarer Identifikationsstrategie;
- **B:** peer-reviewtes, häufig präregistriertes Experiment oder quasi-experimentelle Studie mit klarer Begrenzung;
- **C:** etablierte theoretische oder konzeptionelle Arbeit, die Konstrukte und Mechanismen begründet;
- **D:** offizieller Datenbericht, Umfrage oder noch nicht peer-reviewter Preprint.

Der Grad bezeichnet die Art der Evidenz, nicht automatisch die Richtigkeit jeder Interpretation. Auch große Metaanalysen können hohe Heterogenität und Publikationsbias aufweisen; auch kleine Experimente können für einen eng definierten Mechanismus sehr informativ sein.

3. Begriffe: Qualität ist nicht Vielfalt ist nicht Strategie

3.1 Generative KI

Generative KI bezeichnet Modelle, die neue Inhalte wie Text, Bild, Audio, Video oder Code aus statistisch gelernten Datenmustern erzeugen. Große Sprachmodelle werden in ihrer Grundform darauf trainiert, die bedingte Wahrscheinlichkeit des nächsten Tokens aus dem vorhergehenden Kontext zu schätzen. Das erklärt ihre Fähigkeit, sprachlich plausible Muster zu reproduzieren; es beweist jedoch nicht, dass jedes Ergebnis notwendig „durchschnittlich“ sein muss. Sampling, Modellarchitektur, Ausrichtung, Werkzeuge, Retrieval, Fine-Tuning, Prompting und proprietärer Kontext verändern die Ausgabe erheblich.

Für die Markenführung ist entscheidend, dass ein allgemeines Modell zunächst allgemeine Muster kennt. Es besitzt nicht automatisch die nichtöffentlichen Ursachen einer Marke: ihre tatsächlichen Fähigkeiten, internen Konflikte, Opferentscheidungen, Kundenkenntnis, Produktgrenzen, Belege und zukünftigen Verpflichtungen. Diese Informationen müssen als verifizierte Kontexte, Regeln oder Entscheidungen in den Prozess gelangen.

3.2 Fünf getrennte Bewertungsdimensionen

Ausführungsqualität bezeichnet handwerkliche Güte: Verständlichkeit, sprachliche oder visuelle Stimmigkeit, Fehlerarmut, Realismus, Relevanz zum Briefing und gegebenenfalls messbare kurzfristige Reaktion.

Individuelle Neuartigkeit bezeichnet den Abstand eines einzelnen Ergebnisses zu Erwartungen, üblichen Lösungen oder einem Referenzraum.

Kollektive Vielfalt bezeichnet die Unterschiedlichkeit einer Menge von Ergebnissen untereinander. Eine Menge kann aus lauter gut bewerteten Einzelideen bestehen und dennoch eng, redundant oder homogen sein.

Strategische Eigenständigkeit bezeichnet die Passung eines Ergebnisses zu einer unverwechselbaren Position, einem belastbaren Belegsystem und verbindlichen Ausschlüssen. Sie ist nicht mit Neuartigkeit identisch: Eine Strategie kann vertraute Ausdrucksmittel nutzen und dennoch eigenständig sein, wenn sie eine reale, schwer kopierbare Konfiguration von Fähigkeiten und Entscheidungen sichtbar macht.

Markenwirkung bezeichnet Folgen in Wahrnehmung und Verhalten, etwa Aufmerksamkeit, Erinnerung, erwünschte Assoziationen, Glaubwürdigkeit, Berücksichtigung, Wahl oder Loyalität. Eine hohe Bewertung eines einzelnen Assets ist dafür weder notwendig noch hinreichend.

3.3 Arbeitsdefinition von „Mittelmaß“

Im weiteren Verlauf meint „Mittelmaß“ nicht schlechte Gestaltung. Gemeint ist ein Zustand aus drei Merkmalen:

- **hohe Basiskompetenz:** Der Output ist formal brauchbar und oft professionell;
- **distributionale Zentralität:** Der Output liegt nahe häufigen Mustern und ähnelt anderen Outputs;
- **strategische Austauschbarkeit:** Der Output trägt wenig, das nur diese Marke glaubwürdig sagen, zeigen oder unterlassen kann.

Diese Definition macht die Ausgangsthese empirisch prüfbar. Sie erlaubt zugleich die scheinbar widersprüchliche Aussage, dass KI die durchschnittliche Qualität hebt und den Anteil strategisch austauschbarer Kommunikation dennoch erhöhen kann.

4. Die Produktionsökonomie kippt

4.1 Kosten und Verbreitung

Die Kosten leistungsfähiger Modellnutzung sind in kurzer Zeit drastisch gefallen. Der Stanford AI Index 2025 beziffert die Kosten für eine Modellabfrage auf dem Niveau von GPT-3.5 im MMLU-Benchmark mit 20 US-Dollar je Million Tokens im November 2022 und 0,07 US-Dollar im Oktober 2024 – ein Rückgang um mehr als den Faktor 280. Der AI Index 2026 berichtet, dass 88 Prozent der befragten Organisationen KI einsetzen und 70 Prozent generative KI in mindestens einer Geschäftsfunktion nutzen. Diese Zahlen sind keine global repräsentativen Bevölkerungswerte; sie belegen jedoch, dass die Technologie in der untersuchten Unternehmenspopulation vom Experiment zur verbreiteten Infrastruktur geworden ist.

Sinkende Inferenzkosten sind nicht identisch mit sinkenden Gesamtkosten. Gute Systeme benötigen Datenpflege, Zugriffsrechte, Integration, Qualitätssicherung, Sicherheit, Training und Governance. Trotzdem verändert sich die relative Kostenstruktur: Das Erzeugen einer weiteren Variante wird erheblich billiger als das Beschaffen neuer Marktwahrheiten, das Treffen einer verbindlichen Entscheidung oder das Tragen ihrer Folgen.

4.2 Die Forschung widerspricht der These vom bloßen Qualitätsverfall

Noy und Zhang (2023) randomisierten 453 Hochschulabsolventinnen und -absolventen mit berufsspezifischen, incentivierten Schreibaufgaben. ChatGPT senkte die durchschnittliche Bearbeitungszeit um 40 Prozent und erhöhte die extern bewertete Qualität um 18 Prozent. Die Leistungsunterschiede zwischen Teilnehmenden nahmen ab, weil schwächere Ausgangsleistungen stärker profitierten.

Brynjolfsson, Li und Raymond (2025) untersuchten die gestaffelte Einführung eines KI-Assistenten bei 5.172 Servicekräften und mehr als drei Millionen Chats. Die Zahl gelöster Fälle pro Stunde stieg im Mittel um 15 Prozent. Weniger erfahrene und geringer bewertete Beschäftigte verbesserten Geschwindigkeit und Qualität besonders deutlich; bei den besten Beschäftigten waren die Geschwindigkeitsgewinne klein und die Gesprächsqualität ging leicht zurück. Die Kommunikationsmuster konvergierten, weil weniger leistungsstarke Beschäftigte den Mustern der stärkeren ähnlicher wurden.

Dell'Acqua et al. (2026) führten mit 758 Beraterinnen und Beratern der Boston Consulting Group ein präregistriertes Feldexperiment durch. Bei 18 Aufgaben innerhalb der damaligen Leistungsgrenze von GPT-4 erledigten KI-Nutzende 12,2 Prozent mehr Aufgaben, arbeiteten 25,1 Prozent schneller und erzielten je nach KI-Bedingung im zentralen Qualitätsmaß rund 30 bis 34 Prozent höhere Werte. Die Befunde widerlegen die Vorstellung, KI könne nur billige, schwache Ausführung liefern.

Auch in der visuellen Marketingproduktion ist der Gegenbefund stark. Hartmann, Exner und Domdey (2025) verglichen 10.320 mit sieben Text-zu-Bild-Modellen erzeugte Marketingbilder mit 2.400 menschlich produzierten Bildern und erhoben 254.400 Bewertungen. Die besten synthetischen Bilder übertrafen menschliche Vergleichsbilder bei mehreren Qualitäts-, Realismus- und Ästhetikmaßen. In einem Feldtest mit mehr als 173.000 Impressionen erreichten KI-Banner bis zu 50 Prozent höhere Klickraten als ein Banner mit professionellem Stockfoto.

Diese Studien messen unterschiedliche Aufgaben, Zeiträume und Qualitätskriterien; ihre Prozentwerte sind nicht direkt vergleichbar. Gemeinsam zeigen sie jedoch: **Die wissenschaftlich haltbare Kritik an KI kann nicht darin bestehen, ihre Ausführungsqualität pauschal abzuwerten.**

4.3 Vom knappen Entwurf zur knappen Auswahl

Wenn Varianten teuer sind, begrenzt Produktionskapazität die Kommunikation. Wenn Varianten billig werden, steigt die Zahl möglicher Ausgaben schneller als die menschliche Fähigkeit, sie zu prüfen, einzuordnen und konsistent zu führen. Dieses Muster entspricht einem bekannten Problem der Informationsüberlastung: Sobald Informationsmenge und -komplexität die Verarbeitungskapazität überschreiten, verschlechtern sich Suche, Analyse und Entscheidungen (Eppler & Mengis, 2004; Roetzel, 2019).

Die technologische Produktivität löst den Engpass also nicht ersatzlos auf. Sie verlagert ihn. Drei neue Kosten werden sichtbarer:

- **Selektionskosten:** Welche der vielen brauchbaren Varianten ist für diese Marke richtig?
- **Koordinationskosten:** Bleiben Aussagen, Bilder, Kanäle und Zeitpunkte miteinander vereinbar?
- **Fehlerkosten:** Wie häufig werden plausible, aber falsche, unklare oder markenfremde Inhalte veröffentlicht?

Die OECD (2025) fasst die junge Produktivitätsliteratur ähnlich zusammen: Kurzfristige Leistungsgewinne sind über verschiedene Aufgaben hinweg real, aber heterogen; organisatorische Anpassungen, Prozesse, Fähigkeiten und Kontrolle bestimmen, ob sie sich in nachhaltige Produktivität übersetzen.

Abbildung 1 · Knappheitsverschiebung bei skalierbarer Kommunikation

5. Das Mittelmaß-Paradox: Der Durchschnitt steigt, die Vielfalt kann fallen

5.1 Ein wiederkehrendes Verteilungsmuster

Der entscheidende Befund der Kreativitätsforschung ist kein einfacher Qualitätsverlust. Er ist ein **Verteilungsparadox**: KI kann die Leistung einzelner Personen erhöhen und gleichzeitig die Unterschiedlichkeit der Gesamtmenge reduzieren.

Doshi und Hauser (2024) randomisierten 293 Autorinnen und Autoren in drei Bedingungen: ohne KI, mit einer GPT-4-Idee oder mit bis zu fünf GPT-4-Ideen. 600 weitere Personen bewerteten die Kurzgeschichten. KI-Zugang erhöhte Neuartigkeit, Nützlichkeit, Schreibqualität und Unterhaltung, besonders bei weniger kreativen Schreibenden. Gleichzeitig waren KI-unterstützte Geschichten einander signifikant ähnlicher. Die Autoren beschreiben dies als soziales Dilemma: Individuell verbessert sich das Ergebnis, kollektiv verengt sich der Raum neuartiger Inhalte.

Zhou und Lee (2024) analysierten mehr als vier Millionen Kunstwerke von über 50.000 Nutzerinnen und Nutzern einer Plattform. Nach Einführung von Text-zu-Bild-KI stiegen Produktivität und die als Favoriten pro Ansicht gemessene Bewertung. Zugleich nahm die durchschnittliche inhaltliche und visuelle Neuartigkeit ab. Die maximale inhaltliche Neuartigkeit stieg, während maximale und durchschnittliche visuelle Neuartigkeit sanken. Das ist kein reiner Niedergang: Die kreative Grenze kann sich erweitern, obwohl der durchschnittliche Output stärker konvergiert. Entscheidend wird, wer explorieren und aus vielen Varianten kohärent filtern kann.

Lee und Chung (2024) fanden in fünf Experimenten, dass ChatGPT die durchschnittliche Kreativität von Ideen gegenüber Bedingungen ohne Technologie oder mit Websuche erhöhte, vor allem bei inkrementellen Ideen. Eine Reanalyse derselben Experimente durch Meincke, Nave und Terwiesch (2025) zeigte jedoch, dass die Ideenpools unter ChatGPT-Bedingungen weniger vielfältig waren. Der Widerspruch ist nur scheinbar: Die erste Analyse bewertet die Qualität einzelner Ideen, die zweite die Varianz der Menge.

Moon, Green und Kushlev (2025) untersuchten in drei präregistrierten Studien 2.200 College-Essays. Mit zunehmender Zahl hinzugefügter Texte brachten menschliche Essays mehr neue Ideen in den Pool ein als GPT-4-Essays. Der Unterschied blieb trotz Prompt- und Parameteränderungen bestehen. Die Autoren folgern nicht, dass einzelne KI-Texte unbrauchbar seien, sondern dass breite LLM-Nutzung die kollektive Ideenvielfalt reduzieren kann.

Wenger und Kenett (2026) erweiterten den Test über ein einzelnes Modell hinaus. Sie verglichen 102 Menschen mit 22 Sprachmodellen auf drei standardisierten Kreativitätsaufgaben. Die Modelle erreichten im Mittel ungefähr menschliche Werte individueller Originalität; ihre Antworten waren auf Populationsebene in allen drei Aufgaben signifikant weniger variabel. Kreativere Systemprompts erhöhten die Einzelwerte leicht, beseitigten die Varianzlücke aber nicht. Der Befund spricht dagegen, das Problem allein einem einzelnen Anbieter oder einem ungeschickten Standardprompt zuzuschreiben.

Eine noch nicht peer-reviewte Metaanalyse von Holzner, Maier und Feuerriegel (2025) bündelte 28 Studien mit 8.214 Teilnehmenden. Sie fand keinen signifikanten Unterschied zwischen GenAI und Menschen in individueller kreativer Leistung, einen positiven Effekt der Mensch–KI-Zusammenarbeit gegenüber Menschen allein und einen deutlich negativen Effekt auf die Ideenvielfalt. Wegen ihres Preprint-Status und der Heterogenität der eingeschlossenen Aufgaben ist die konkrete Effektgröße vorläufig; die Richtung stimmt jedoch mit mehreren peer-reviewten Einzelstudien überein.

5.2 Warum „mehr Ideen“ nicht automatisch mehr Vielfalt bedeutet

Fluency, also die Zahl erzeugter Ideen, und Flexibilität beziehungsweise Vielfalt sind unterschiedliche Kreativitätsmerkmale. Ein System kann tausend Varianten aus einem engen semantischen Gebiet liefern. Der Eindruck großer Auswahl entsteht dann aus Oberfläche, Formulierung oder Stil, während die zugrunde liegenden Konzepte wiederkehren.

Für Organisationen ist das besonders tückisch, weil Evaluationsprozesse oft lokal funktionieren. Eine einzelne KI-Idee wird gegen eine leere Seite oder einen menschlichen Erstentwurf verglichen und wirkt stark. Selten wird sie gegen die gleichzeitig erzeugten Ideen tausender anderer Marken verglichen. Das individuelle Produktivitätsmaß belohnt die erste Verbesserung; der Markt bestraft später die kollektive Ähnlichkeit.

Die Markenrelevanz folgt daraus nicht automatisch. Die Kreativitätsstudien untersuchen Kurzgeschichten, Essays, Kunst oder Laboraufgaben, nicht langfristige Markenführung. Der Transfer ist dennoch plausibel, weil Markenkommunikation ebenfalls aus wiederholten symbolischen Entscheidungen in einem gemeinsamen Aufmerksamkeitsraum besteht. Er bleibt eine begründete Generalisierung und muss künftig direkt getestet werden.

5.3 Das Risiko ist bedingt, nicht unvermeidlich

Homogenisierung ist keine Naturkonstante jedes KI-Einsatzes. Modelle können mit höherer Temperatur, alternativen Suchverfahren, vielfältigen Rollen, adversarial prompting, getrennten Ideationssträngen, proprietären Daten, Retrieval und menschlichen Ausgangsideen breitere Räume erkunden. Meincke, Mollick und Terwiesch (2024) zeigen in einem Working Paper, dass gezielte Promptstrategien die Varianz von KI-Ideen deutlich erhöhen können. Wenger und Kenett (2026) zeigen zugleich, dass einfache Aufforderungen zu „mehr Kreativität“ die Lücke zur menschlichen Vielfalt nicht schließen.

Die praktische Konsequenz lautet daher nicht, KI aus Kreativprozessen zu entfernen. Sie lautet, Vielfalt **zu entwerfen und zu messen**, statt sie aus der Zahl der Varianten abzuleiten. Ein Team sollte getrennt prüfen:

- Wie gut ist die beste Einzelidee?
- Wie groß ist der konzeptionelle Abstand zwischen den Ideen?
- Wie viele Ideen beruhen auf unterschiedlichen Wahrheiten statt auf umformulierten Mustern?
- Welche Idee kann diese Marke glaubwürdig besitzen?
- Welche Idee würde ohne Markennamen trotzdem eindeutig zugeordnet?

6. Die gezackte Leistungsgrenze: Produktion ist nicht Urteil

6.1 Gute Form kann falsche Entscheidungen verdecken

Das BCG-Experiment von Dell'Acqua et al. (2026) enthält neben den starken Produktivitätsgewinnen einen für Markenführung zentralen Gegenbefund. Bei einer komplexen Aufgabe außerhalb der damaligen Leistungsgrenze von GPT-4 lösten 84,5 Prozent der Kontrollgruppe die Aufgabe korrekt, aber nur 60 beziehungsweise 70,6 Prozent der beiden KI-Gruppen. Im Mittel waren KI-Nutzende 19 Prozentpunkte seltener korrekt. Gleichwohl wirkten auch falsche KI-unterstützte Antworten subjektiv kohärenter und überzeugender.

Dieser Befund trennt zwei Fähigkeiten, die im Arbeitsalltag leicht verwechselt werden: eine Antwort professionell formulieren und erkennen, ob das zugrunde liegende Problem richtig gelöst ist. Für Marken entsteht daraus ein spezifisches Risiko. Je besser ein System falsche Positionierungen, unbelegte Versprechen oder unpassende Zielgruppenentscheidungen formuliert, desto weniger sichtbar ist der Fehler im Entwurf selbst.

Die Studie hat eine wichtige Grenze: Die negative Aussage beruht auf einer einzigen Aufgabe außerhalb der Leistungsgrenze. Sie beweist nicht, dass KI generell schlechte strategische Entscheidungen trifft. Sie zeigt aber, dass die Leistungsgrenze „gezackt" und für Nutzende schwer erkennbar sein kann. Gute Ergebnisse in benachbarten Aufgaben sind keine Garantie für gutes Urteil in der nächsten.

6.2 Mensch–KI ist nicht automatisch besser als das jeweils beste System

Vaccaro, Almaatouq und Malone (2024) werteten in einer präregistrierten Metaanalyse 106 Experimente mit 370 Effektgrößen aus. Mensch–KI-Kombinationen waren im Mittel deutlich besser als Menschen allein. Sie waren jedoch schlechter als der jeweils bessere Einzelakteur aus Mensch oder KI. Besonders bei Entscheidungsaufgaben zeigten sich Verluste; bei Inhaltserzeugung waren die Gewinne größer. Die Heterogenität war sehr hoch, und der Suchzeitraum endete im Juni 2023. Das Ergebnis ist deshalb keine pauschale Regel für heutige Modelle. Es widerlegt aber die bequeme Annahme, dass „human in the loop" bereits durch Anwesenheit eines Menschen Qualität garantiert.

Ein Mensch kann KI-Fehler übernehmen, sich von flüssiger Sprache ankern lassen oder eine plausible Empfehlung mangels eigener Sicherheit nicht verwerfen. Verantwortliche Kontrolle erfordert daher Wissen, Zeit, Kriterien und echte Entscheidungsbefugnis.

6.3 Kritisches Denken verlagert sich

Lee et al. (2025) befragten 319 Wissensarbeiterinnen und Wissensarbeiter zu 936 konkreten GenAI-Nutzungsfällen. Höheres Vertrauen in GenAI war mit weniger selbst berichtetem kritischem Denken verbunden; höheres Vertrauen in die eigene Aufgabenkompetenz mit mehr. Die qualitative Analyse zeigte eine Verschiebung von originärer Erzeugung zu Verifikation, Integration und Prozessverantwortung.

Die Studie ist korrelativ und basiert auf Selbstauskünften. Sie belegt keine kausale Erosion kritischer Fähigkeiten. Für Markenprozesse liefert sie dennoch eine plausible Aufgabenbeschreibung: Die anspruchsvolle Arbeit verschwindet nicht, sondern wandert in die Prüfung von Quellen, Annahmen, Konsequenzen und Passung.

Raisch und Krakowski (2021) beschreiben diesen Zusammenhang als Automations–Augmentations-Paradox. Automatisierung und Ergänzung lassen sich organisatorisch nicht sauber trennen; übermäßiger Fokus auf eine Seite erzeugt neue Abhängigkeiten. Die strategische Frage lautet deshalb nicht, ob Menschen oder KI „führen“, sondern welche Entscheidungen codifiziert, welche unterstützt und welche unter menschlicher Verantwortung gehalten werden müssen.

7. Warum Marken von Auswahl, Ausschluss und Wiederholung leben

7.1 Markenwert entsteht nicht aus Outputmenge

Keller (1993) definiert kundenbasierten Markenwert als die unterschiedliche Reaktion von Konsumentinnen und Konsumenten auf Marketing aufgrund ihres Markenwissens. Dieses Wissen besteht aus Bekanntheit und einem Netzwerk von Assoziationen. Starke Marken benötigen nach dieser Logik starke, günstige und – für die vorliegende These entscheidend – einzigartige Assoziationen.

Die Zahl veröffentlichter Inhalte ist darin kein eigener Werttreiber. Inhalte sind Mittel, um Gedächtnisstrukturen zu bilden, zu bestätigen oder zu verändern. Werden sie austauschbar, widersprüchlich oder nicht zugeordnet, können auch professionell produzierte Assets wenig Markenwert aufbauen.

Keller und Lehmann (2006) beschreiben Positionierung als Richtungsvorgabe für Marketingaktivitäten: Sie legt fest, was eine Marke in ihrem Marketing tun und nicht tun sollte. Park, Jaworski und MacInnis (1986) behandeln den Markenkern als über Zeit zu wählendes, einzuführendes, auszuarbeitendes und zu festigendes Konzept. Urde (1999) erweitert diese Sicht organisatorisch: Markenorientierung bedeutet, Identität nicht als Kommunikationsdekoration, sondern als strategische Ressource zu entwickeln und zu schützen.

7.2 Glaubwürdigkeit ist ein kumulatives Signal

In der Informationsökonomie reduzieren Marken Unsicherheit, wenn sie glaubwürdige Signale über Position und Leistungsfähigkeit senden. Erdem und Swait (1998, 2004) zeigen, dass Glaubwürdigkeit – verstanden als wahrgenommene Vertrauenswürdigkeit und Kompetenz – die Aufnahme einer Marke in das Consideration Set und die Wahlwahrscheinlichkeit erhöht. Konsistenz, Klarheit und vergangene Investitionen tragen zur Glaubwürdigkeit eines Signals bei.

Erdem, Swait und Valenzuela (2006) validierten das Signalmodell mit Befragungs- und Experimental-daten zu Orangensaft und Computern in Brasilien, Deutschland, Indien, Japan, Spanien, der Türkei und den USA. Marken fungierten länderübergreifend als Signale von Produktpositionen; der Nutzen glaubwürdiger Marken war bei hoher Unsicherheitsvermeidung und kollektivistischer Orientierung größer.

Für KI-gestützte Kommunikation folgt daraus eine Asymmetrie. Eine weitere Aussage ist billig. Ein glaubwürdiges Signal ist es nicht, weil Glaubwürdigkeit aus überprüfbarer Fähigkeit, konsistentem Verhalten und der Bereitschaft entsteht, sich an ein Versprechen binden zu lassen. KI kann das Signal formulieren; sie kann die zugrunde liegende Verpflichtung nicht anstelle der Organisation erfüllen.

7.3 Strategie entsteht durch Trade-offs

Porter (1996) trennt operative Effektivität von Strategie. Produktivitätswerkzeuge machen viele Unternehmen gleichzeitig schneller; ein dauerhafter Unterschied entsteht erst durch eine eigenständige Position und ein zusammenpassendes Aktivitätssystem. Sein prägnanter Satz lautet: „The essence of strategy is choosing what not to do."

Für generative Systeme ist Ausschluss besonders wichtig. Ein offener Prompt maximiert den Möglichkeitsraum. Eine Marke muss ihn begrenzen. Sie braucht Antworten auf Fragen, die ein allgemeines Modell nicht normativ entscheiden kann:

- Welche Zielgruppe erhält Vorrang, obwohl andere erreichbar wären?
- Welches Kundenproblem wird bewusst nicht gelöst?
- Welches Preis-, Qualitäts- oder Serviceniveau ist verbindlich?
- Welche Belege reichen für eine Aussage – und welche nicht?
- Welche Tonalität passt selbst dann nicht, wenn sie kurzfristig Reichweite bringt?
- Welche gesellschaftliche oder kulturelle Position ist aus der Realität der Organisation ableitbar, und welche wäre opportunistische Pose?
- Welche Technologien nutzt die Marke intern, aber nicht in der sichtbaren Leistung?

Solche Ausschlüsse wirken zunächst wie verlorene Optionen. Strategisch erzeugen sie jedoch Kohärenz, erleichtern Auswahl und machen Erwartungen prüfbar. Gerade weil KI nahezu jede Position sprachlich plausibel darstellen kann, gewinnt die reale Bindung an eine Position an Bedeutung.

7.4 Differenzierung und Distinktheit sind nicht dasselbe

Marken müssen nicht in jeder Produkteigenschaft einzigartig sein. Romaniuk, Sharp und Ehrenberg (2007) zeigten über 17 Kategorien in Australien und Großbritannien, dass die Mehrheit der Käufer ihre genutzte Marke nicht ausdrücklich als anders oder einzigartig wahrnahm. Die Autoren verschieben den Fokus deshalb von behaupteter Differenzierung auf **Distinktheit**: gelernte, eindeutige Signale, die eine Marke erkennen und erinnern lassen.

Diese Korrektur ist für KI-Zeiten wesentlich. Strategische Eigenständigkeit und perceptual distinctiveness ergänzen sich:

- **Strategische Eigenständigkeit** beantwortet, welche Bedeutung, Fähigkeit und Verpflichtung die Marke besitzt.
- **Distinktheit** beantwortet, durch welche wiederholten verbalen, visuellen und auditiven Codes sie schnell erkannt wird.

KI kann fortlaufend neue Stile erzeugen. Markenführung muss entscheiden, welche wenigen Codes nicht fortlaufend ersetzt werden. Ohne diese Disziplin kann Variantenreichtum die Wiedererkennbarkeit schwächen.

Pieters, Warlop und Wedel (2002) zeigen zudem, dass Originalität in Werbung Aufmerksamkeit fördern kann, während Vertrautheit die Markenverarbeitung unterstützt. Smith et al. (2007) modellieren Werbekreativität als Zusammenspiel von Divergenz und Relevanz. Neuheit ohne Markenbezug kann Aufmerksamkeit vom Absender abziehen; Relevanz ohne Divergenz kann im Umfeld untergehen. Gute Markenkommunikation benötigt beides – und eine konsistente Zuordnung.

7.5 Aufmerksamkeit wird durch mehr Output nicht vermehrt

Die Literatur zu Informationsüberlastung dokumentiert seit Jahrzehnten, dass steigende Informationsmenge bei begrenzter Verarbeitungskapazität zu suboptimaler Suche und Entscheidung

führen kann (Eppler & Mengis, 2004; Roetzel, 2019). Daraus folgt nicht automatisch, dass jede zusätzliche Markenbotschaft schädlich ist. Wohl aber steigt die Beweislast für ihren Beitrag.

Wenn alle Marktteilnehmer ihre Frequenz erhöhen, kann der relative Aufmerksamkeitsgewinn einzelner Marken gering bleiben, während Prüf-, Media- und Koordinationskosten wachsen. Der Produktionsvorteil wird so teilweise zu einem Red-Race-Effekt: Jede Marke produziert mehr, ohne dass der Markt mehr erinnern kann.

8. Authentizität: Nicht jeder KI-Einsatz ist kommunikativ neutral

Brüns und Meißner (2024) untersuchten in drei Experimenten Reaktionen auf den Einsatz generativer KI in Social-Media-Inhalten von Marken. Wahrgenommene KI-Nutzung löste negativere Einstellungen und Verhaltensabsichten aus; der Effekt wurde durch geringere wahrgenommene Markenauthentizität vermittelt. Er fiel schwächer aus, wenn KI Menschen unterstützte, statt sie vollständig zu ersetzen.

Kirk und Givi (2025) zeigten in sieben präregistrierten Experimenten einen ähnlichen, aber kontextabhängigen „AI-authorship effect“. Wenn emotionale Markenkommunikation als KI-verfasst galt, sanken positive Mundpropaganda und Loyalität; wahrgenommene Inauthentizität und moralischer Ekel vermittelten den Effekt. Bei faktischen Botschaften, reiner KI-Bearbeitung oder direkter Signatur durch die KI war er schwächer. Der Befund ist daher kein allgemeines Verbot von KI-Autorschaft. Er zeigt, dass die erwartete menschliche Quelle Teil der Botschaft sein kann.

Dem steht die Leistungsstudie von Hartmann et al. (2025) gegenüber: KI-generierte visuelle Assets können kurzfristige Marketingkennzahlen klar übertreffen. Beide Befundgruppen sind vereinbar. Ein anonymes oder funktionales Banner wird nach anderen Kriterien bewertet als eine emotionale Erklärung, ein Trauertext, eine Entschuldigung oder ein Bekenntnis zur Herkunft. Der Wert menschlicher Urheberschaft ist aufgaben- und erwartungsabhängig.

Für Marken entsteht damit eine weitere strategische Entscheidung: nicht nur **wie** KI eingesetzt wird, sondern **wo** ihre Nutzung zur sichtbaren Bedeutung der Kommunikation gehört. Transparenz darf weder als pauschale Vertrauensgarantie noch als pauschale Strafe behandelt werden. Sie muss zum Kommunikationsgegenstand, zur Erwartung des Publikums und zur tatsächlichen Rollenverteilung passen.

9. Die strategische Selektionsprämie

9.1 Definition

Dieses Paper bezeichnet als **strategische Selektionsprämie** den zunehmenden Hebel einer verbindlichen Auswahlentscheidung in einem System stark gesunkener Generationskosten.

Die Prämie ist kein bereits gemessener Preisaufschlag und keine Bilanzgröße. Sie ist ein konzeptioneller Mechanismus. Eine strategische Entscheidung – etwa eine Positionierung, ein Belegstandard oder ein Ausschluss – wirkt nicht nur auf einen einzelnen Inhalt, sondern auf eine große Zahl zukünftiger Inhalte. Sinkt der Preis jedes einzelnen Outputs und steigt deren Zahl, steigt die Reichweite der vorgelagerten Entscheidung.

Vereinfacht:

Erwarteter Entscheidungshebel = Zahl der gesteuerten Outputs × Wirkungsdifferenz zwischen passender und unpassender Auswahl × Wert beziehungsweise Schaden je Wirkungseinheit.

Diese Darstellung ist keine empirisch geschätzte Gleichung. Sie macht die Logik prüfbar. Wenn eine Markenregel 500 Ausgaben steuert statt fünf, vervielfacht sich ihr potenzieller Nutzen – und ebenso ihr Fehler.

9.2 Fünf abgeleitete Propositionen

Proposition 1: Der Qualitätsboden steigt schneller als die Unterscheidbarkeit

Generative KI hebt bei vielen Aufgaben die Mindestqualität und komprimiert Leistungsunterschiede. Dadurch wird formale Professionalität weniger selten. Ein Wettbewerber kann ähnliche handwerkliche Qualität mit demselben Basismodell, vergleichbaren Prompts und niedrigen Kosten erreichen. Der Vorteil wandert zu Input, Auswahl, System und realer Leistung.

Proposition 2: Gemeinsame Modelle erzeugen ein Konvergenzrisiko auf Populationsebene

Die Evidenz aus Geschichten, Essays, Kunst, Brainstorming und standardisierten Kreativitätstests zeigt wiederholt: Einzelwerte können steigen, während die Menge homogener wird. Für Marken ist das Risiko am größten, wenn Modelle aus generischen Briefings ohne proprietäre Tatsachen, alternative Suchpfade oder Diversitätskontrolle erzeugen.

Proposition 3: Der Ausschlusswert wächst mit dem Möglichkeitsraum

Je mehr plausible Optionen verfügbar sind, desto wichtiger werden Regeln, die Optionen verwerfen. „Was wir nicht tun“ ist keine negative Ergänzung der Positionierung, sondern eine operative Steuerung des Generationsraums. Ausschlüsse sparen Prüfung, erhöhen Konsistenz und reduzieren opportunistische Botschaften.

Proposition 4: Proprietäre Wahrheit wird zum wertvollsten Prompt

Allgemeine Modelle teilen allgemeines Wissen. Schwerer kopierbar sind verifizierte Kundenerkenntnisse, eigene Daten, Produktfähigkeiten, Herkunft, Prozesse, Beziehungen, Fachwissen und konkrete Entscheidungen. Diese Inhalte sind nicht deshalb wertvoll, weil sie menschlich sind, sondern weil sie real, spezifisch und zurechenbar sind.

Proposition 5: Verantwortete Auswahl ist ein Markenasset

Ein Modell kann Optionen bewerten, aber keine institutionelle Verantwortung tragen. Markenentscheidungen binden Organisationen über Zeit, Budgets, Verhalten und Erwartungen. Wer eine Aussage freigibt, entscheidet nicht nur über Text, sondern über ein Versprechen. Diese Verantwortungsfähigkeit bleibt ein knappes Komplement zur automatisierten Produktion.

9.3 Warum der Effekt ökonomisch plausibel ist

Die Forschung zu General-Purpose-Technologien zeigt, dass Produktivitätsgewinne komplementäre immaterielle Investitionen erfordern. Brynjolfsson, Rock und Syverson (2021) nennen unter anderem neue Prozesse, Fähigkeiten und organisatorisches Kapital. Teece (1986) argumentiert allgemeiner, dass der Ertrag einer Innovation von komplementären Assets abhängt, nicht nur von der Erfindung selbst.

Übertragen auf generative KI bedeutet das: Zugang zum Modell ist leicht imitierbar. Schwieriger imitierbar sind eine klare Markenkonstitution, gepflegte Wissensbestände, distinctive assets,

Freigaberegeln, tatsächliche Fähigkeiten und ein lernendes Evaluationssystem. Der strategische Vorteil liegt nicht im Prompt als Textfragment, sondern im organisatorischen System, das festlegt, welche Wahrheiten in den Prompt gelangen und welche Ergebnisse den Markt erreichen.

10. Das Vaternam-Modell: Markenkonstitution, Exploration, Gate

Das aus der Evidenz abgeleitete Modell trennt drei Ebenen, die in vielen KI-Workflows vermischt werden.

10.1 Ebene I: Die Markenkonstitution

Die Markenkonstitution ist kein Tonalitäts-PDF. Sie enthält überprüfbare und entscheidungsfähige Grundlagen:

- **Realität:** Was kann die Organisation nachweislich leisten?
- **Adressat:** Für wen wird vorrangig Wert geschaffen?
- **Position:** Welches Problem, welche Kategorie oder welche Erwartung wird besetzt?
- **Versprechen:** Welche Wirkung darf kommuniziert werden?
- **Belege:** Welche Evidenz autorisiert welche Aussage?
- **Trade-offs:** Was wird trotz möglicher Nachfrage nicht angeboten, behauptet oder bedient?
- **Codes:** Welche verbalen, visuellen und auditiven Signale werden über Zeit wiederholt?
- **Grenzen:** Welche rechtlichen, ethischen, kulturellen und reputativen Risiken führen zum Ausschluss?

Diese Ebene bleibt menschlich verantwortlich, kann aber KI-gestützt analysiert werden. KI darf Widersprüche finden, Optionen simulieren und Gegenargumente erzeugen. Die Organisation muss entscheiden, welche Option wahr, tragfähig und bindend ist.

10.2 Ebene II: Exploration und Produktion

Innerhalb der Konstitution kann KI ihre Stärke ausspielen:

- breite Ideenräume erkunden;
- Varianten für Zielgruppen und Kanäle erzeugen;
- komplexe Inhalte vereinfachen;
- Formate übertragen und lokalisieren;
- alternative Frames und Gegenpositionen testen;
- visuelle und sprachliche Prototypen beschleunigen;
- große Mengen konsistent auf bekannte Regeln anwenden.

Exploration sollte bewusst diversifiziert werden. Statt einen Prompt zwanzigmal zu wiederholen, können Teams voneinander unabhängige Suchpfade definieren: funktional, emotional, kulturell, kategorial, kontraintuitiv, kundenzentriert, beweiszentriert oder wettbewerbsadversarial. Erst danach werden Ergebnisse zusammengeführt.

10.3 Ebene III: Das redaktionelle Gate

Das Gate entscheidet, was die Marke tatsächlich sagt. Es benötigt Kriterien, nicht nur Geschmack. Ein belastbares Gate prüft mindestens:

5. **Wahrheit:** Ist die Aussage faktisch belegt und aktuell?
6. **Besitzfähigkeit:** Kann die Marke diese Aussage glaubwürdig besetzen?

7. **Ausschluss:** Welche attraktive Alternative wird durch die Wahl verworfen?
8. **Zuordnung:** Stärkt der Inhalt erkennbare Markencodes oder nur das Format?
9. **Konsequenz:** Welche Erwartung oder Verpflichtung erzeugt die Veröffentlichung?
10. **Portfoliofit:** Ergänzt der Inhalt das bestehende Kommunikationssystem oder dupliziert er es?
11. **Varianz:** Fügt er dem Markt eine neue Perspektive hinzu oder nur eine neue Formulierung?
12. **Aufgabengrenze:** Ist KI hier Generator, Bearbeiter, Entscheider oder lediglich Werkzeug – und ist diese Rolle angemessen?

Das Gate kann teilautomatisiert vorprüfen, etwa Quellen, verbotene Begriffe, Inkonsistenzen oder Ähnlichkeit. Die Freigabe strategischer, emotionaler oder haftungsrelevanter Kommunikation bleibt zurechenbar.

10.4 Der Lernkreislauf

Leistungsdaten fließen zurück, aber nicht ungefiltert. Kurzfristige Klicks dürfen langfristige Markenwirkung nicht allein definieren. Das System trennt mindestens:

- operative Signale: Zeit, Kosten, Fehler, Conversion;
- kommunikative Signale: Aufmerksamkeit, Verständnis, Erinnerung, Zuordnung;
- strategische Signale: gewünschte Assoziationen, Glaubwürdigkeit, Consideration, Preis- oder Präferenzwirkung;
- negative Signale: Verwechslung, Überversprechen, Authentizitätsverlust, regulatorisches oder reputatives Risiko.

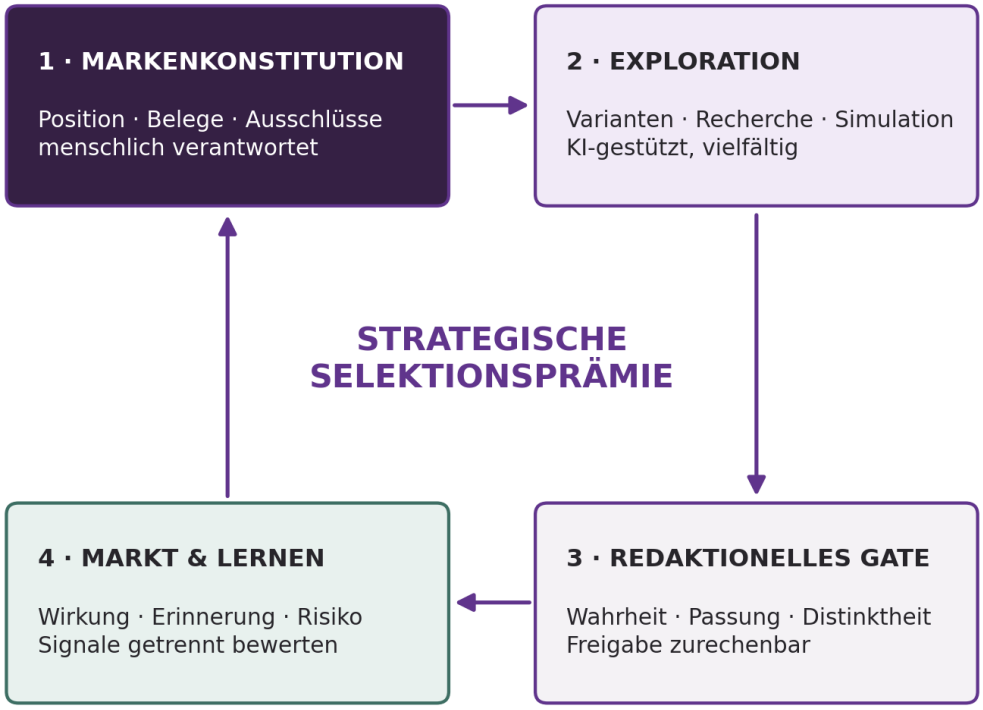

Abbildung 2 · Vaternam-Entscheidungssystem für KI-gestützte Markenführung

11. Sechs Tests gegen strategische Austauschbarkeit

11.1 Der Wettbewerbertest

Wenn ein Wettbewerber Name, Farbe und Produkt austauschen könnte, ohne dass der Inhalt unplausibel würde, ist die Ausführung möglicherweise gut, aber die Aussage strategisch schwach. Der Test ist streng: Auch Kategorien mit ähnlichen Leistungen benötigen keine künstliche Einzigartigkeit, wohl aber eine erkennbare Auswahl und eindeutige Zuordnung.

11.2 Der Belegtest

Jede zentrale Behauptung erhält einen Belegtyp: Messwert, Zertifikat, Prozess, Kundenbeispiel, beobachtbares Verhalten, Vertrag, Garantie oder nachvollziehbare Herkunft. Fehlt der Beleg, wird die Aussage als Hypothese oder Ambition gekennzeichnet – nicht als Tatsache veröffentlicht.

11.3 Der Opfertest

Eine Position ohne Kosten ist häufig nur eine Präferenzformulierung. Der Opfertest fragt: Welche Umsatzchance, Zielgruppe, Funktion, Tonalität oder kurzfristige Reichweite nimmt die Organisation nicht mit, um die Position glaubwürdig zu halten? Nicht jede Marke braucht dramatische Opfer. Aber jede belastbare Position braucht Grenzen.

11.4 Der Gedächtnistest

Welche zwei bis vier Signale sollen nach wiederholtem Kontakt erinnert und der Marke zugeordnet werden? Jede neue Variante wird danach bewertet, ob sie diese Gedächtnisstruktur stärkt oder unnötig ersetzt. Neuheit ist kein Selbstzweck, wenn sie den Absender löscht.

11.5 Der Portfoliotest

Ein einzelner Inhalt kann stark und das Gesamtportfolio trotzdem redundant sein. Der Portfoliotest misst Themen-, Argument-, Bild- und Formatähnlichkeit über Zeit. Er entdeckt, ob KI Varianten produziert, die lediglich dieselbe Idee anders verpacken.

11.6 Der Verantwortlichkeitstest

Vor Veröffentlichung wird eine Person oder Funktion benannt, die für Wahrheit, Markenpassung und Konsequenzen einsteht. „Die KI hat es formuliert“ ist weder Quelle noch Verantwortungsmodell.

12. Was Marken jetzt anders bewerten sollten

12.1 Produktivität und Markenwert getrennt messen

Zeitersparnis, Stückkosten und Outputmenge sind Produktivitätsmetriken. Sie sagen wenig darüber aus, ob Markenwissen aufgebaut wurde. Ein KI-Programm kann operativ erfolgreich und strategisch neutral oder schädlich sein. Controlling sollte deshalb beide Ebenen getrennt ausweisen.

12.2 Variantenbudgets in Entscheidungsbudgets umschichten

Wenn Produktion billiger wird, ist es rational, einen Teil der Einsparung in bessere Briefings, Marktforschung, proprietäre Daten, Expertenprüfung, Pretests und Portfolioanalyse zu investieren. Die falsche Reaktion wäre, jede eingesparte Produktionsminute ausschließlich in mehr Output umzuwandeln.

12.3 Seniorität an den Engpass verlagern

Erfahrene Markenverantwortliche werden nicht dadurch wertvoll, dass sie jeden Erstentwurf selbst schreiben. Ihr größter Hebel liegt in Problemdefinition, Trade-offs, Belegstandards, Bewertung und Verantwortung. Das bedeutet nicht, Juniorarbeit zu entwerten; im Gegenteil kann KI Ausführungskompetenz verbreitern. Organisationen müssen jedoch verhindern, dass Nachwuchskräfte nur noch auswählen, ohne eigene Urteilsfähigkeit aufzubauen.

12.4 Markenwissen maschinenlesbar, aber nicht beliebig machen

Styleguides sollten zu gepflegten Wissens- und Regelsystemen werden: mit Quellen, Gültigkeitsdaten, Anspruchsklassen, Beispielen und klaren Ausschlüssen. Retrieval und Fine-Tuning können Generik reduzieren. Technische Anpassung ersetzt jedoch nicht die vorgelagerte Entscheidung, welche Inhalte wahr und strategisch bindend sind.

12.5 KI-Rollen nach Kommunikationsaufgabe unterscheiden

Faktische Produktbeschreibung, emotionales Gründerstatement, Krisenkommunikation, Performance-Banner und langfristige Markenpositionierung stellen unterschiedliche Erwartungen an Urheberschaft, Belege und Verantwortung. Eine einzige unternehmensweite Regel „KI erlaubt“ oder „KI verboten“ ist deshalb zu grob.

13. Was dieses Paper nicht behauptet

13.1 Es behauptet nicht, dass Menschen grundsätzlich kreativer sind

Mehrere Studien zeigen, dass KI auf standardisierten Kreativitätsmaßen menschliche Durchschnittswerte erreicht oder übertrifft. Die stärksten Menschen können weiterhin überlegen sein; auch das ist aufgabenspezifisch. Der relevante Befund ist die Kombination aus hoher Einzelqualität und geringerer Populationsvarianz.

13.2 Es behauptet nicht, dass Homogenisierung unvermeidlich ist

Prompts, Modelle, Daten, Sampling, menschliche Ideen und organisatorische Verfahren können Vielfalt erhöhen. Das Paper behauptet nur, dass Vielfalt nicht aus hoher Outputmenge folgt und in mehreren Studien ohne gezielte Gegenmaßnahmen sank.

13.3 Es setzt kreative Homogenisierung nicht mit „Model Collapse“ gleich

Shumailov et al. (2024) zeigen, dass rekursives Training auf modellgenerierten Daten unter bestimmten Bedingungen zu Verteilungsverlusten und Modellkollaps führen kann. Gerstgrasser et al. (2024) zeigen, dass das Beibehalten und Akkumulieren realer Daten diesen Effekt in ihren Experimenten vermeiden kann. Das sind Befunde über Trainingsdaten über Modellgenerationen hinweg. Sie beweisen nicht, dass

ein aktueller KI-Entwurf für eine Marke mittelmäßig ist. Die beiden Phänomene teilen das Thema Varianzverlust, dürfen aber nicht verwechselt werden.

13.4 Es behauptet keinen direkt gemessenen Preis für Strategie

Die strategische Selektionsprämie ist eine theoretische Synthese. Um sie kausal zu messen, wären Feldexperimente nötig, die Unternehmen mit identischem KI-Zugang, aber unterschiedlich klaren Markenkonstitutionen über längere Zeit vergleichen und neben Produktivität auch Markenassoziationen, Zuordnung, Consideration, Preiswirkung und Glaubwürdigkeit messen.

14. Forschungsagenda

Die Forschung zur Markenführung mit generativer KI steht am Anfang. Besonders relevant sind fünf offene Fragen.

Erstens: Wie verändert KI-gestützte Massenproduktion die semantische Distanz zwischen Wettbewerberkommunikationen innerhalb realer Kategorien? Dafür wären große, längsschnittliche Text-, Bild- und Video-Korpora vor und nach GenAI-Adoption nötig.

Zweitens: Welche Markenkontexte reduzieren Homogenisierung am wirksamsten – proprietäre Daten, negative Regeln, Beispiele, Fine-Tuning, Retrieval oder menschliche Ausgangsideen?

Drittens: Wann führt Konsistenz zu Markenaufbau und wann zu Langeweile? Distinktheit verlangt Wiederholung; Kreativität verlangt relevante Abweichung. KI erlaubt, diese Grenze experimentell über viele Varianten zu untersuchen.

Viertens: Wie wirken die Offenlegung der Autorschaft und sichtbare KI-Nutzung in verschiedenen Kulturen, Kategorien und Beziehungssituationen? Die vorhandenen Befunde sind kontextabhängig und nicht als universelle Konsumentenreaktion zu behandeln.

Fünftens: Welche Fähigkeiten erhalten menschliches Urteil? Langzeitstudien sollten prüfen, ob kontinuierliche KI-Unterstützung Expertise aufbaut, verlagert oder abbaut – und welche Trainings- und Freigabedesigns diese Effekte beeinflussen.

15. Schlussfolgerung

Die Ausgangsthese muss verschärft werden, indem man sie korrigiert. Generative KI produziert nicht einfach unendlich viel schlechte Kommunikation. Sie produziert sehr viel brauchbare, oft gute und gelegentlich herausragende Kommunikation zu schnell fallenden Grenzkosten. Genau darin liegt die strategische Herausforderung.

Wenn Professionalität verbreitet wird, unterscheidet Professionalität allein weniger. Wenn Ideen massenhaft erzeugt werden, wird nicht die Idee knapp, sondern eine begründete Auswahl. Wenn jedes Versprechen in Sekunden formulierbar ist, gewinnt der Beleg. Wenn jeder Stil simulierbar ist, gewinnt Wiedererkennbarkeit. Wenn jede Haltung sprachlich darstellbar ist, gewinnt die reale Verpflichtung. Und wenn jede Option verfügbar bleibt, gewinnt der Mut zum Ausschluss.

Der wirtschaftliche Wert verschiebt sich deshalb nicht von Menschen zu Maschinen oder von Maschinen zu Menschen. Er verschiebt sich innerhalb des Systems: von der Produktion einzelner Ausgaben zur Architektur der Entscheidungen, die viele Ausgaben steuert.

KI verbilligt das Erzeugen. Sie verteuert nicht automatisch die Strategie – aber sie vergrößert deren Hebel.

Anhang A: Evidenzmatrix

Tabelle A1 · Empirische Evidenz zu Produktivität, Qualität und Vielfalt

Studie · Evidenzgrad	Design	Zentraler Befund	Begrenzung
Noy & Zhang (2023) · B	RCT; n=453; berufliche Schreibaufgaben	–40 % Bearbeitungszeit; +18 % bewertete Qualität; stärkere Gewinne bei schwächeren Ausgangsleistungen	Kurze Aufgaben; frühe ChatGPT-Version; keine Langzeitwirkung
Brynjolfsson et al. (2025) · A	Gestaffelte Einführung; n=5.172; >3 Mio. Servicechats	+15 % gelöste Fälle je Stunde; größte Gewinne bei weniger Erfahrenen; Kommunikationsmuster konvergieren	Ein Unternehmen und Aufgabentyp; firmenspezifisches Assistenzsystem
Dell'Acqua et al. (2026) · B	Präregistriertes RCT; n=758 BCG-Berater	Innerhalb der Frontier: +12,2 % Aufgaben, –25,1 % Zeit, ca. +30–34 % Qualität; außerhalb: 19 Prozentpunkte weniger korrekte Lösungen	Außerhalb-Frontier-Befund beruht auf einer Aufgabe; GPT-4, Stand 2023
Vaccaro et al. (2024) · A	Metaanalyse; 106 Experimente, 370 Effekte, 74 Arbeiten	Mensch+KI besser als Mensch allein (g=0,64), aber schlechter als das jeweils beste Teammitglied (g=–0,23); Content-Erzeugung günstiger als Entscheidungen	Hohe Heterogenität; Literatursuche bis Juni 2023
Doshi & Hauser (2024) · B	Präregistriert; 293 Schreibende, 600 Bewertende	KI-Ideen erhöhen Einzelbewertungen; Geschichten werden einander ähnlicher	Kurze fiktionale Texte; definierte Kreativitätsmaße
Zhou & Lee (2024) · B	>4 Mio. Kunstwerke; >50.000 Nutzer, >5.800 Adopter	Produktivität +25 %; Favoriten je Ansicht +50 %; mittlere Neuartigkeit sinkt, maximale Inhaltsneuartigkeit steigt	Plattformkontext; Adoption nicht vollständig randomisiert
Moon et al. (2025) · B	Drei präregistrierte Studien; 2.200 College-Essays	Jeder zusätzliche menschliche Text erweitert den Ideenraum stärker als ein zusätzlicher GPT-4-Text	Essayaufgabe; Ergebnisse nicht automatisch auf alle Medien übertragbar
Wenger & Kenett (2026) · B	102 Menschen, 22 LLMs; drei Kreativitätsaufgaben	Ähnliche mittlere Originalität; LLMs zeigen in allen Aufgaben geringere Populationsvarianz	Standardisierte Aufgaben; Modelle und Systemprompts verändern sich
Hartmann et al. (2025) · B	10.320 KI- vs. 2.400 Humanbilder; 254.400 Ratings; >173.000 Impressions	Beste KI-Bilder übertreffen Vergleichsbilder in mehreren Ratings; im Feld bis zu +50 % CTR gegenüber Stockfoto	Kurzfristige visuelle und Performance-Maße; kein langfristiger Markenaufbau

Legende: A = Metaanalyse oder große Feldstudie mit belastbarer Identifikationsstrategie; B = peer-reviewtes Experiment oder quasi-experimentelle Studie; C = etablierte theoretische Arbeit.

Tabelle A2 · Evidenzbasis für den Transfer zur Markenführung

Studie · Evidenzgrad	Design	Zentraler Befund	Begrenzung
Keller (1993) · C	Konzeption und Messlogik der Customer-Based Brand Equity	Markenwissen beruht auf starken, günstigen und einzigartigen Assoziationen	Grundlagentheorie; nicht GenAI-spezifisch
Erdem & Swait (1998, 2004) · B/C	Signaling-Modell und empirische Wahlmodelle	Glaubwürdige Marken senken wahrgenommenes Risiko und Informationskosten und erhöhen Berücksichtigung/Wahl	Transfer auf KI-Kommunikation ist theoretisch
Erdem et al. (2006) · B	Cross-Country-Validierung in sieben Ländern	Marken wirken kulturübergreifend als Signale; Stärke variiert mit Kulturmerkmalen	Nicht alle Länder/Kategorien; Daten vor GenAI
Pieters et al. (2002) · B	Werbe- und Aufmerksamkeitsforschung	Originalität und Vertrautheit beeinflussen Markenaufmerksamkeit und Erinnerung	Werbemittelkontext; kein allgemeines Markenwertmaß
Romaniuk et al. (2007) · B	17 Märkte in Australien und Großbritannien	Wenige Käufer beschreiben Marken als unterschiedlich; Distinktheit und eindeutige Zuordnung bleiben zentral	Zwei Länder; Differenzierung und Distinktheit sind getrennte Konstrukte

Studie · Evidenzgrad	Design	Zentraler Befund	Begrenzung
Smith et al. (2007) · B	Modellierung kreativer Werbung	Werbekreativität entsteht aus dem Zusammenspiel von Divergenz und Relevanz und beeinflusst Verarbeitung und Reaktion	Werbekontext; Kreativität allein garantiert keinen Markenaufbau
Brüns & Meißner (2024) · B	Drei Experimente zu GenAI-Content	Wahrgenommener KI-Ersatz kann Authentizität und Reaktionen mindern; Assistenz wird günstiger bewertet als Ersatz	Reaktionen sind auf Format, Disclosure und Kontext angewiesen
Kirk & Givi (2025) · B	Sieben präregistrierte Experimente	Bei emotionalen Botschaften kann vermutete KI-Autorenschaft Loyalität und positive Mundpropaganda über Inauthentizität/moralischen Ekel senken	Effekt schwächer bei Fakten, KI-Editing oder sichtbarer menschlicher Autorenschaft

Legende: A = Metaanalyse oder große Feldstudie mit belastbarer Identifikationsstrategie; B = peer-reviewtes Experiment oder quasi-experimentelle Studie; C = etablierte theoretische Arbeit.

Anhang B: Such- und Prüfprotokoll

B.1 Suchfelder und Beispielbegriffe

Die Recherche kombinierte unter anderem folgende englische und deutsche Begriffe:

- “generative AI productivity experiment”, “LLM knowledge work quality”, “human AI synergy meta-analysis”;
- “generative AI creativity diversity”, “LLM homogenization”, “idea diversity brainstorming”, “collective novelty”;
- “generative AI marketing content authenticity”, “AI authorship brand loyalty”, “AI advertising effectiveness”;
- “brand equity unique associations”, “brand credibility signal”, “brand distinctiveness”, “strategic positioning trade-offs”;
- “information overload decision making”, “attention advertising originality familiarity”;
- “model collapse synthetic data real data”, “AI complementary intangible assets”.

B.2 Prüfregeln

Quantitative Angaben wurden, soweit öffentlich zugänglich, mit Abstract, Volltext oder offizieller Metadatenquelle abgeglichen. DOI, Jahr, Journal und Stichprobe wurden gegen Verlags-, PubMed-, RePEc- oder Institutionsseiten geprüft. Preprints werden als solche bezeichnet. Prozentwerte aus unterschiedlichen Studien werden nicht als direkt vergleichbare Effektgrößen behandelt. Kausalsprache wird nur bei randomisierten oder geeigneten quasi-experimentellen Designs verwendet.

B.3 Ausschlüsse

Nicht aufgenommen wurden nicht nachvollziehbare Marketingstatistiken, anonyme Umfragen ohne Methodennachweis, reine Anbieterbehauptungen, Sekundärberichte bei verfügbarem Primärbeleg sowie Arbeiten, deren Bezug zur Forschungsfrage nur über mehrere spekulative Zwischenschritte herstellbar wäre.

B.4 Transparenz zum KI-Einsatz

Generative KI wurde zur Rechercheunterstützung, Strukturierung, sprachlichen Überarbeitung und Dokumentproduktion eingesetzt. Die Verantwortung für Auswahl, Interpretation, Quellenprüfung und die in diesem Paper vorgeschlagene Synthese liegt bei Vaternam Research. Die Verwendung eines generativen Werkzeugs ist kein Beleg; als Belege dienen ausschließlich die ausgewiesenen Quellen.

Literatur

Brüns, J. D., & Meißner, M. (2024). Do you create your content yourself? Using generative artificial intelligence for social media content creation diminishes perceived brand authenticity. *Journal of Retailing and Consumer Services, 79*, 103790. https://doi.org/10.1016/j.jretconser.2024.103790

Boussioux, L., Lane, J. N., Zhang, M., Jacimovic, V., & Lakhani, K. R. (2024). The crowdless future? Generative AI and creative problem-solving. *Organization Science, 35*(5), 1589–1607. https://doi.org/10.1287/orsc.2023.18430

Brynjolfsson, E., Li, D., & Raymond, L. R. (2025). Generative AI at work. *The Quarterly Journal of Economics, 140*(2), 889–942. https://doi.org/10.1093/qje/qjae044

Brynjolfsson, E., Rock, D., & Syverson, C. (2021). The productivity J-curve: How intangibles complement general purpose technologies. *American Economic Journal: Macroeconomics, 13*(1), 333–372. https://doi.org/10.1257/mac.20180386

Dell'Acqua, F., McFowland, E., III, Mollick, E., Lifshitz, H., Kellogg, K. C., Rajendran, S., Krayer, L., Candelon, F., & Lakhani, K. R. (2026). Navigating the jagged technological frontier: Field experimental evidence of the effects of artificial intelligence on knowledge worker productivity and quality. *Organization Science, 37*(2), 403–423. https://doi.org/10.1287/orsc.2025.21838

Doshi, A. R., & Hauser, O. P. (2024). Generative AI enhances individual creativity but reduces the collective diversity of novel content. *Science Advances, 10*(28), eadn5290. https://doi.org/10.1126/sciadv.adn5290

Eppler, M. J., & Mengis, J. (2004). The concept of information overload: A review of literature from organization science, accounting, marketing, MIS, and related disciplines. *The Information Society, 20*(5), 325–344. https://doi.org/10.1080/01972240490507974

Erdem, T., & Swait, J. (1998). Brand equity as a signaling phenomenon. *Journal of Consumer Psychology, 7*(2), 131–157. https://doi.org/10.1207/S15327663JCP0702_02

Erdem, T., & Swait, J. (2004). Brand credibility, brand consideration, and choice. *Journal of Consumer Research, 31*(1), 191–198. https://doi.org/10.1086/383434

Erdem, T., Swait, J., & Valenzuela, A. (2006). Brands as signals: A cross-country validation study. *Journal of Marketing, 70*(1), 34–49. https://doi.org/10.1509/jmkg.70.1.034.qxd

Gerstgrasser, M., Schaeffer, R., Dey, A., Rafailov, R., Pai, D., Sleight, H., Hughes, J., Korbak, T., Agrawal, R., Gromov, A., Roberts, D. A., Yang, D., Donoho, D. L., & Koyejo, S. (2024). Is model collapse inevitable? Breaking the curse of recursion by accumulating real and synthetic data [Preprint]. *arXiv*. https://doi.org/10.48550/arXiv.2404.01413

Hartmann, J., Exner, Y., & Domdey, S. (2025). The power of generative marketing: Can generative AI create superhuman visual marketing content? *International Journal of Research in Marketing, 42*(1), 13–31. https://doi.org/10.1016/j.ijresmar.2024.09.002

Holzner, N., Maier, S., & Feuerriegel, S. (2025). Generative AI and creativity: A systematic literature review and meta-analysis [Preprint]. *arXiv*. https://doi.org/10.48550/arXiv.2505.17241

Huang, M.-H., & Rust, R. T. (2025). The GenAI future of consumer research. *Journal of Consumer Research, 52*(1), 4–17. https://doi.org/10.1093/jcr/ucaf013

Keller, K. L. (1993). Conceptualizing, measuring, and managing customer-based brand equity. *Journal of Marketing, 57*(1), 1–22. https://doi.org/10.1177/002224299305700101

Keller, K. L., & Lehmann, D. R. (2006). Brands and branding: Research findings and future priorities. *Marketing Science, 25*(6), 740–759. https://doi.org/10.1287/mksc.1050.0153

Kirk, C. P., & Givi, J. (2025). The AI-authorship effect: Understanding authenticity, moral disgust, and consumer responses to AI-generated marketing communications. *Journal of Business Research, 186*, 114984. https://doi.org/10.1016/j.jbusres.2024.114984

Lee, B. C., & Chung, J. (2024). An empirical investigation of the impact of ChatGPT on creativity. *Nature Human Behaviour, 8*(10), 1906–1914. https://doi.org/10.1038/s41562-024-01953-1

Lee, H.-P., Sarkar, A., Tankelevitch, L., Drosos, I., Rintel, S., Banks, R., & Wilson, N. (2025). The impact of generative AI on critical thinking: Self-reported reductions in cognitive effort and confidence effects from a survey of knowledge workers. In *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems* (Article 1121, pp. 1–22). ACM. https://doi.org/10.1145/3706598.3713778

Meincke, L., Mollick, E., & Terwiesch, C. (2024). Prompting diverse ideas: Increasing AI idea variance [Preprint]. *arXiv*. https://doi.org/10.48550/arXiv.2402.01727

Meincke, L., Nave, G., & Terwiesch, C. (2025). ChatGPT decreases idea diversity in brainstorming. *Nature Human Behaviour, 9*(6), 1107–1109. https://doi.org/10.1038/s41562-025-02173-x

Moon, K., Green, A. E., & Kushlev, K. (2025). Homogenizing effect of large language models (LLMs) on creative diversity: An empirical comparison of human and ChatGPT writing. *Computers in Human Behavior: Artificial Humans, 6*, 100207. https://doi.org/10.1016/j.chbah.2025.100207

Noy, S., & Zhang, W. (2023). Experimental evidence on the productivity effects of generative artificial intelligence. *Science, 381*(6654), 187–192. https://doi.org/10.1126/science.adh2586

OECD. (2025). *The effects of generative AI on productivity, innovation and entrepreneurship*. OECD Artificial Intelligence Papers. https://doi.org/10.1787/b21df222-en

Park, C. W., Jaworski, B. J., & MacInnis, D. J. (1986). Strategic brand concept-image management. *Journal of Marketing, 50*(4), 135–145. https://doi.org/10.1177/002224298605000401

Pieters, R., Warlop, L., & Wedel, M. (2002). Breaking through the clutter: Benefits of advertisement originality and familiarity for brand attention and memory. *Management Science, 48*(6), 765–781. https://doi.org/10.1287/mnsc.48.6.765.192

Porter, M. E. (1996). What is strategy? *Harvard Business Review, 74*(6), 61–78. https://hbr.org/1996/11/what-is-strategy

Raisch, S., & Krakowski, S. (2021). Artificial intelligence and management: The automation–augmentation paradox. *Academy of Management Review, 46*(1), 192–210. https://doi.org/10.5465/amr.2018.0072

Roetzel, P. G. (2019). Information overload in the information age: A review of the literature from business administration, business psychology, and related disciplines with a bibliometric approach and framework development. *Business Research, 12*(2), 479–522. https://doi.org/10.1007/s40685-018-0069-z

Romaniuk, J., Sharp, B., & Ehrenberg, A. (2007). Evidence concerning the importance of perceived brand differentiation. *Australasian Marketing Journal, 15*(2), 42–54. https://doi.org/10.1016/S1441-3582(07)70042-3

Shumailov, I., Shumaylov, Z., Zhao, Y., Gal, Y., Papernot, N., & Anderson, R. (2024). AI models collapse when trained on recursively generated data. *Nature, 631*, 755–759. https://doi.org/10.1038/s41586-024-07566-y

Smith, R. E., MacKenzie, S. B., Yang, X., Buchholz, L. M., & Darley, W. K. (2007). Modeling the determinants and effects of creativity in advertising. *Marketing Science, 26*(6), 819–833. https://doi.org/10.1287/mksc.1070.0272

Stanford Institute for Human-Centered Artificial Intelligence. (2025). *AI Index Report 2025*. Stanford University. https://hai.stanford.edu/ai-index/2025-ai-index-report

Stanford Institute for Human-Centered Artificial Intelligence. (2026). *AI Index Report 2026*. Stanford University. https://hai.stanford.edu/ai-index/2026-ai-index-report

Teece, D. J. (1986). Profiting from technological innovation: Implications for integration, collaboration, licensing and public policy. *Research Policy, 15*(6), 285–305. https://doi.org/10.1016/0048-7333(86)90027-2

Urde, M. (1999). Brand orientation: A mindset for building brands into strategic resources. *Journal of Marketing Management, 15*(1–3), 117–133. https://doi.org/10.1362/026725799784870504

Vaccaro, M., Almaatouq, A., & Malone, T. (2024). When combinations of humans and AI are useful: A systematic review and meta-analysis. *Nature Human Behaviour, 8*, 2293–2303. https://doi.org/10.1038/s41562-024-02024-1

Wenger, E., & Kenett, Y. N. (2026). Large language models are homogeneously creative. *PNAS Nexus, 5*(3), pgag042. https://doi.org/10.1093/pnasnexus/pgag042

Zhou, E., & Lee, D. (2024). Generative artificial intelligence, human creativity, and art. *PNAS Nexus, 3*(3), pgae052. https://doi.org/10.1093/pnasnexus/pgae052

Zitierempfehlung

Vaternam Research. (2026). *Das Mittelmaß-Paradox: Generative KI hebt die Qualität – und macht strategische Markenentscheidungen wertvoller* (Vaternam Research Paper, Version 1.0). Vaternam Kreativagentur e.K. https://www.vaternam.de/