IA GENERATIVA NA COMUNICAÇÃO GERENCIAL: EVIDÊNCIAS EXPERIMENTAIS SOBRE O MODELO DE CONFIRMAÇÃO DE EXPECTATIVAS

GENERATIVE AI IN MANAGERIAL COMMUNICATION: EXPERIMENTAL EVIDENCE ON THE EXPECTATION-CONFIRMATION MODEL

LEONARDO DE SOUZA PANDOLFINI
UNINOVE – UNIVERSIDADE NOVE DE JULHO

FERNANDO ANTONIO RIBEIRO SERRA
UNINOVE – UNIVERSIDADE NOVE DE JULHO

Comunicação:
O XIV SINGEP foi realizado em conjunto com a 14th Conferência Internacional do CIK (CYRUS Institute of Knowledge) e com a SPARK (Society for Promoting Academic Research and Knowledge), em formato híbrido, com sede presencial na UNINOVE - Universidade Nove de Julho, no Brasil.

IA GENERATIVA NA COMUNICAÇÃO GERENCIAL: EVIDÊNCIAS EXPERIMENTAIS SOBRE O MODELO DE CONFIRMAÇÃO DE EXPECTATIVAS

Objetivo do estudo
Investigar, com base no Modelo de Confirmação de Expectativas, se o uso do ChatGPT em uma tarefa realista de comunicação gerencial melhora o desempenho e a satisfação de profissionais de gerenciamento de projetos, mediado pela revisão da utilidade percebida após a adoção.

Relevância/originalidade
A pesquisa oferece evidências experimentais sobre os efeitos causais da IA generativa no desempenho e na satisfação, fortalece o Modelo de Confirmação de Expectativas e orienta gestores, PMOs e entidades certificadoras na adoção estruturada, verificação crítica e governança responsável dessas ferramentas digitais.

Metodologia/abordagem
Realizou-se experimento randomizado entre 64 profissionais de projetos, distribuídos igualmente entre controle e ChatGPT, que redigiram comunicação executiva. Compararam-se tempo, qualidade, produtividade percebida, satisfação e intenção de continuidade mediante testes t, avaliações especializadas e mediação por regressão com bootstrap de 1.000 reamostras.

Principais resultados
Um experimento com 64 profissionais demonstrou que o ChatGPT reduziu em 73% o tempo de elaboração de relatórios e elevou sua qualidade em quatro critérios. A produtividade percebida mediou metade do efeito sobre a satisfação, oferecendo evidências experimentais à gestão de projetos.

Contribuições teóricas/metodológicas
A pesquisa oferece evidências experimentais sobre os efeitos causais da IA generativa no desempenho e na satisfação, fortalece o Modelo de Confirmação de Expectativas e orienta gestores, PMOs e entidades certificadoras na adoção estruturada, verificação crítica e governança responsável dessas ferramentas digitais.

Contribuições sociais/para a gestão
A contribuição social reside em demonstrar como a IA generativa pode reduzir o esforço dedicado a tarefas administrativas e melhorar a clareza das comunicações, liberando gestores para decisões, negociação e relacionamento com partes interessadas, desde que seu uso inclua supervisão humana responsável.

Palavras-chave: Inteligencia artificial generativa, ChatGPT, Gestao de Projetos, Comunicacao gerencial, Expectation-Confirmation Model

GENERATIVE AI IN MANAGERIAL COMMUNICATION: EXPERIMENTAL EVIDENCE ON THE EXPECTATION-CONFIRMATION MODEL

Study purpose
To investigate, based on the Expectation-Confirmation Model, whether using ChatGPT in a realistic managerial communication task improves project management professionals' performance and satisfaction, mediated by the post-adoption reassessment of perceived usefulness.

Relevance / originality
The study provides experimental evidence on the causal effects of generative AI on performance and satisfaction, strengthens the Expectation-Confirmation Model, and guides managers, PMOs, and certification bodies in the structured adoption, critical evaluation, and responsible governance of these digital tools.

Methodology / approach
A randomized experiment was conducted with 64 project management professionals, equally assigned to control and ChatGPT groups, who drafted an executive communication Completion time, quality, perceived productivity, satisfaction, and continuance intention were compared using t-tests, expert ratings, and regression-based mediation with 1,000

Main results
An experiment involving 64 professionals demonstrated that ChatGPT reduced report completion time by 73% and improved quality across four criteria. Perceived productivity mediated half of the effect on satisfaction, providing experimental evidence for the field of project management.

Theoretical / methodological contributions
The study provides experimental evidence on the causal effects of generative AI on performance and satisfaction, strengthens the Expectation-Confirmation Model, and guides managers, PMOs, and certification bodies in the structured adoption, critical evaluation, and responsible governance of these digital tools.

Social / management contributions
The social contribution lies in demonstrating how generative AI can reduce the effort devoted to administrative tasks and improve communication clarity, freeing managers to focus on decision-making, negotiation, and stakeholder engagement, provided that its use includes responsible human oversight.

Keywords: Generative artificial intelligence, ChatGPT, Project Management, Managerial communication, Expectation-Confirmation Model

IA GENERATIVA NA COMUNICAÇÃO GERENCIAL: EVIDÊNCIAS EXPERIMENTAIS SOBRE O MODELO DE CONFIRMAÇÃO DE EXPECTATIVAS

1 Introdução

A comunicação é uma das atividades mais críticas da gestão de projetos, tanto para a coordenação das partes interessadas quanto pelo tempo gerencial que consome (Rabechini Jr. & Carvalho, 2003; Zhou & Gorman, 2024). Relatórios executivos de status, briefings para patrocinadores e mensagens formais exigem interpretação de dados, síntese textual e adaptação a diferentes públicos, geralmente sob pressão de tempo. A execução eficiente dessas tarefas é relevante, pois falhas de comunicação estão associadas a retrabalho, atrasos e desalinhamento entre partes interessadas (Zerfass, Hagelstein & Tench, 2020; Qosidah & Susilo, 2024).

A difusão da IA generativa, exemplificada pelo ChatGPT, introduziu novas possibilidades para o trabalho intensivo em texto (Brown et al., 2020; Eloundou, Manning, Mishkin & Rock, 2024). Experimentos em áreas correlatas identificaram ganhos de produtividade e qualidade, embora sua magnitude dependa do alinhamento entre as capacidades da ferramenta e as exigências da tarefa (Noy & Zhang, 2023; Peng, Kalliamvakou, Cihon & Demirer, 2023; Dell'Acqua et al., 2023; Brynjolfsson, Li & Raymond, 2025). Na gestão de projetos, entretanto, a literatura concentra-se principalmente em levantamentos sobre adoção e impactos percebidos (Müller, Locatelli, Holzmann, Nilsson & Sagay, 2024; Felicetti, Cimino, Mazzoleni & Ammirato, 2024; Dwivedi et al., 2025), sem examinar experimentalmente os efeitos de um episódio específico de uso em uma tarefa concreta.

Essa lacuna também possui uma dimensão teórica. Para compreender se a IA generativa apoia a comunicação gerencial, não basta investigar a disposição inicial para adotá-la; é necessário analisar as avaliações formadas após o uso. O Modelo de Confirmação de Expectativas — ECM — (Bhattacherjee, 2001) propõe que a experiência pós-uso modifica a utilidade percebida, influenciando a satisfação e a intenção de continuidade. Embora o modelo tenha sido aplicado à IA generativa em estudos transversais com estudantes e consumidores (Kavitha & Joshith, 2025; Tian et al., 2024), seu mecanismo ainda não foi testado experimentalmente com gestores de projetos em uma tarefa realista. Assim, este estudo investiga se o uso da IA generativa produz os resultados objetivos e avaliativos previstos pelo ECM e se seu efeito sobre a satisfação é mediado pela produtividade percebida.

Para responder a essa questão, realizou-se um experimento controlado com 64 profissionais de gestão de projetos, alocados aleatoriamente entre duas condições. O grupo de controle elaborou uma comunicação executiva utilizando métodos tradicionais, enquanto o grupo experimental realizou a mesma tarefa com acesso ao ChatGPT. Foram examinados o tempo de conclusão, a qualidade avaliada por especialistas independentes, a produtividade percebida, a satisfação e a intenção de continuidade. Os resultados revelaram diferenças substanciais nos desfechos e um efeito indireto significativo sobre a satisfação, mediado pela produtividade percebida, correspondente a aproximadamente metade do efeito total.

O estudo contribui para a literatura ao deslocar a análise da IA generativa em gestão de projetos dos levantamentos agregados para evidências experimentais no nível da tarefa. Também amplia a aplicação do ECM ao testar causalmente, em forma reduzida, sua estrutura mediacional no contexto profissional. Na prática, os achados indicam que a IA generativa pode produzir ganhos relevantes na comunicação gerencial, desde que o gestor permaneça responsável pela validação das interpretações e informações produzidas pela ferramenta.

2 Referencial Teórico

2.1 A comunicação gerencial como tarefa cognitivamente exigente

Relatórios executivos de status, briefings para patrocinadores e mensagens às partes interessadas são atividades cognitivamente intensas e relevantes para a gestão de projetos. Exigem interpretação de indicadores, síntese de informações e adaptação textual sob pressão de tempo, enquanto falhas em sua elaboração podem gerar retrabalho, atrasos e desalinhamento. Essas comunicações também seguem práticas, artefatos e vocabulários internacionalmente padronizados por corpos de conhecimento e sistemas de certificação (PMI, 2021). Sua elaboração requer integrar indicadores quantitativos, como os índices de desempenho de prazo e custo, à interpretação qualitativa dos desvios e das ações necessárias. Diante das assimetrias de conhecimento entre gestores e patrocinadores (Zwikael, Salmona, Meredith & Zarghami, 2023), a eficácia da comunicação depende sobretudo de clareza, objetividade, assertividade e adequação ao público.

2.2 A inteligência artificial generativa no trabalho do conhecimento

Os grandes modelos de linguagem representam uma classe qualitativamente nova de ferramentas para o trabalho do conhecimento, produzindo linguagem natural fluente, sintetizando insumos heterogêneos e adaptando os resultados a diferentes públicos (Brown et al., 2020; Eloundou, Manning, Mishkin & Rock, 2024). Experimentos controlados em domínios do trabalho do conhecimento convergem para um padrão consistente: o acesso a um modelo generativo reduz o tempo de conclusão enquanto mantém ou melhora a qualidade do produto, com os maiores ganhos entre trabalhadores menos experientes e menor dispersão de desempenho (Noy & Zhang, 2023; Peng, Kalliamvakou, Cihon & Demirer, 2023; Brynjolfsson, Li & Raymond, 2025). O esforço é comprimido no tempo, a variância é reduzida entre os usuários e a contribuição cognitiva do trabalhador desloca-se da síntese para o julgamento e a verificação (Korinek, 2023).

Esse quadro não é uniforme entre as tarefas. Evidências de uma fronteira tecnológica irregular (jagged technological frontier) (Dell'Acqua et al., 2023) mostram que a IA generativa aprimora o desempenho em tarefas dentro de seus limites de capacidade, mas pode deteriorá-lo quando utilizada além deles. Tarefas que combinam raciocínio numérico com produção narrativa, como a comunicação gerencial ancorada no IDP e no IDC, situam-se em uma zona ambígua dessa fronteira, onde ganhos são possíveis, mas não garantidos. Os ganhos de produtividade documentados em domínios adjacentes, portanto, não podem ser assumidos como transferíveis à comunicação gerencial sem teste empírico.

No âmbito específico da gestão de projetos, a literatura empírica sobre IA generativa cresceu rapidamente, mas permanece ancorada em levantamentos agregados de atitudes de adoção. Evidências em larga escala indicam que os gestores de projetos esperam que a IA transforme o campo, embora a adoção organizacional ainda esteja em estágios iniciais e seja limitada por barreiras de competências e de gestão da mudança (Müller, Locatelli, Holzmann, Nilsson & Sagay, 2024). Pesquisas sobre a apropriação de chatbots generativos por gestores de projetos em exercício identificam a compatibilidade tarefa-tecnologia, as atitudes de inovação e a influência dos pares como fatores que moldam a incorporação do ChatGPT às rotinas profissionais (Felicetti, Cimino, Mazzoleni & Ammirato, 2024), e revisões mais amplas enfatizam confiança, transparência e uso responsável como preocupações emergentes para a integração contínua (Dwivedi et al., 2025; Leavitt, Schabram, Hariharan & Barnes, 2024). O que esses estudos não oferecem e não podem oferecer por design são evidências causais sobre se um episódio específico de uso de IA generativa em uma tarefa realista de gestão de projetos

produz os julgamentos avaliativos que sustentam o uso continuado. Isso requer uma teoria do comportamento pós-adoção e um delineamento que teste seu mecanismo.

2.3 O Modelo de Confirmação de Expectativas e a lacuna abordada por este estudo

O Modelo de Aceitação de Tecnologia (Davis, 1989; Davis, Bagozzi & Warshaw, 1989) e suas extensões (Venkatesh, Morris, Davis & Davis, 2003) explicam a aceitação de tecnologia no momento da primeira exposição por meio de crenças pré-adoção, como utilidade percebida e facilidade de uso percebida. Esses modelos não foram concebidos para explicar por que muitas tecnologias, inicialmente aceitas, são posteriormente abandonadas. O Modelo de Confirmação de Expectativas (Bhattacherjee, 2001), adaptado da teoria de confirmação de expectativas do comportamento do consumidor (Oliver, 1980), foi proposto para preencher essa lacuna, deslocando o foco teórico da aceitação para a continuidade de uso.

O núcleo explicativo do ECM é uma estrutura mediacional específica. A experiência pós-uso molda a utilidade percebida, que é revisada para cima ou para baixo em relação às crenças pré-uso; a utilidade percebida e a confirmação determinam conjuntamente a satisfação; a satisfação e a utilidade percebida, por sua vez, predizem a intenção de continuar usando a tecnologia (Bhattacherjee, 2001; Bhattacherjee & Premkumar, 2004). A satisfação não é mero reflexo da exposição ou das características da tecnologia; ela emerge da avaliação cognitiva da própria experiência. Essa estrutura demonstrou ser robusta em domínios que vão do comércio eletrônico a sistemas corporativos (Lee, 2010), e é o que distingue o ECM de explicações mais simples de efeito direto sobre a resposta à tecnologia.

Duas características da IA generativa tornam o mecanismo do ECM particularmente relevante. Os ganhos de produtividade são experienciados imediatamente, com rascunhos sendo gerados em segundos; essa proximidade temporal comprime o processo de avaliação e torna a experiência de produtividade inusitadamente vívida em comparação às tecnologias tradicionais (Dell'Acqua et al., 2023). A interação também carrega uma dimensão experiencial que vai além da utilidade instrumental, com fluência, senso de competência e alívio cognitivo gerando respostas afetivas que sistemas utilitários tipicamente não produzem (Van der Heijden, 2004; Wang, Rau & Yuan, 2023).

Estudos recentes começaram a aplicar o ECM à IA generativa em contextos educacionais e de consumo (Kavitha & Joshith, 2025; Ngo, Tran & Nguyen, 2024; Tian et al., 2024). Esse conjunto de trabalhos estabelece o ECM como um referencial produtivo para a IA generativa, mas é quase inteiramente transversal: a estrutura mediacional é testada por meio de covariâncias entre constructos em dados de levantamento, e não por meio de comparação controlada entre usuários que experienciaram ou não a tecnologia em uma tarefa realista. Uma busca estruturada na Web of Science e no Scopus em janeiro de 2026 não retornou nenhum estudo anterior que utilizasse atribuição aleatória para testar a lógica mediacional do ECM com usuários profissionais em uma tarefa de trabalho realista. A lacuna abordada pelo presente estudo é, portanto, específica: testar experimentalmente, com profissionais de gestão de projetos, se a estrutura mediacional postulada pelo ECM opera no contexto da IA generativa aplicada à comunicação gerencial.

Duas características do presente delineamento delimitam o escopo do teste que se segue. O ECM clássico inclui a confirmação, ou seja, o grau em que a experiência pós-uso corresponde às expectativas prévias como antecedente da utilidade percebida revisada. Em um experimento entre grupos em que as expectativas pré-tarefa não são elicitadas nem tornadas teoricamente salientes, a confirmação é absorvida pela manipulação da exposição, em vez de ser medida como um constructo separado. O presente estudo fornece, portanto, um teste de uma forma reduzida do ECM, centrado no caminho da exposição à utilidade percebida (operacionalizada

como produtividade percebida), à satisfação e à intenção de continuidade. Os constructos avaliativos são mensurados com itens sintéticos, seguindo a prática comum em estudos experimentais de uso de tecnologia (Wanous, Reichers & Hudy, 1997; Noy & Zhang, 2023).

3 Metodologia

Desenvolvimento das Hipóteses

3.1 O desempenho objetivo como base experiencial

Os ganhos de produtividade documentados em domínios adjacentes do trabalho do conhecimento suscitam a expectativa de que a comunicação gerencial também se beneficiaria da IA generativa, embora tenhamos observado anteriormente que a tarefa de relatório de status situa-se na zona ambígua da fronteira irregular, por combinar raciocínio numérico (IDP, IDC) com síntese narrativa. Duas características desta tarefa específica sugerem, contudo, que os ganhos devem emergir. A parcela narrativa de um relatório de status (enquadramento, adaptação de registro, construção de sentenças) enquadra-se inteiramente dentro da fronteira de capacidade dos modelos atuais. Os elementos estruturais, incluindo cabeçalhos padronizados, formulações padrão e recomendações padrão para desvios de IDP/IDC, são altamente modelados e prontamente produzidos por modelos treinados em comunicações empresariais. A fluência desenvolvida por gestores experientes pode comprimir o tempo de conclusão sem assistência, mas a vantagem mecânica da geração modelada opera em cada execução, enquanto a fluência varia entre os usuários; portanto, a primeira deve predominar em média:

H1. Gestores de projetos que utilizam o ChatGPT concluem a tarefa de comunicação gerencial em menos tempo do que aqueles que utilizam métodos tradicionais.

O tempo de conclusão isolado é um indicador ambíguo: um rascunho mais rápido pode sacrificar o conteúdo analítico. O ECM coloca a utilidade percebida em seu núcleo, mas a utilidade é, em última instância, julgada pela capacidade da comunicação de habilitar o patrocinador a agir. Esse julgamento depende de propriedades reconhecíveis pelos leitores: clareza, objetividade, assertividade e calibração ao público (Zerfass, Hagelstein & Tench, 2020; Zwikael, Salmona, Meredith & Zarghami, 2023). Se os ganhos de velocidade não forem acompanhados de ganhos correspondentes de qualidade, a base experiencial para a atualização das crenças sobre utilidade é enfraquecida e o mecanismo downstream do ECM não deve operar. Avaliadores especialistas independentes fornecem o julgamento de qualidade:

H2. As comunicações executivas produzidas com o suporte do ChatGPT recebem avaliações de qualidade superiores (clareza, objetividade, assertividade e adequação ao público) às produzidas sem ele, conforme julgamento de avaliadores especialistas independentes.

3.2 Os constructos avaliativos do ECM

O desempenho objetivo estabelece a base experiencial, mas não aciona o mecanismo do ECM. O ECM é uma teoria da avaliação pós-uso: descreve como os usuários formam julgamentos avaliativos que sustentam ou desencorajam o uso continuado (Bhattacherjee, 2001). Na formulação clássica, a atualização da utilidade percebida é impulsionada pela confirmação. Como o presente experimento não elicia expectativas pré-tarefa, a confirmação é absorvida pela manipulação da exposição, e as predições teoricamente relevantes tornam-se contrastes entre grupos nos três constructos que o ECM identifica como centrais: utilidade percebida, satisfação e intenção de continuidade.

Em uma tarefa de composição cognitivamente exigente, o componente mais saliente da utilidade é a produtividade percebida, ou seja, a percepção subjetiva do usuário sobre o quanto

o método o auxiliou a produzir o resultado exigido. A IA generativa produz resultados visíveis em segundos, comprimindo o intervalo entre esforço e resultado e tornando a experiência de produtividade inusitadamente vívida (Dell'Acqua et al., 2023):
H3a. Gestores de projetos que utilizam o ChatGPT relatam maior produtividade percebida na tarefa de comunicação do que aqueles que utilizam métodos tradicionais.
A satisfação é conceitualmente distinta da utilidade no ECM. No caso da IA generativa, a imediatez do resultado remove o atrito de uma tarefa intrinsecamente trabalhosa, e a percepção de competência ampliada possibilitada pela ferramenta pode traduzir-se em afeto positivo em relação à experiência:
H3b. Gestores de projetos que utilizam o ChatGPT relatam maior satisfação com a ferramenta ou método utilizado para realizar a tarefa de comunicação do que aqueles que utilizam métodos tradicionais.
A intenção de continuidade é o desfecho comportamental que o ECM identifica como ponto final da avaliação pós-adoção. O ECM prediz que ela decorre do efeito conjunto da utilidade percebida e da satisfação. No presente contexto de sessão única, a medida de continuidade captura a disposição declarada, e não a persistência observada:
H3c. Gestores de projetos que utilizam o ChatGPT relatam maior intenção de utilizar a ferramenta em trabalho real de projetos do que aqueles que utilizam métodos tradicionais.

3.3 O mecanismo mediacional

Efeitos na mesma direção sobre os três constructos avaliativos (H3a, H3b, H3c) são compatíveis com diferentes explicações. Uma explicação de efeito direto poderia prever que a exposição a uma tecnologia em evidência eleva os três constructos em paralelo, por vias não relacionadas à avaliação cognitiva descrita pelo ECM. Uma explicação de aceitação simples poderia prever que os constructos respondem porque o ChatGPT é uma novidade, e não porque os usuários revisaram suas crenças de utilidade com base na experiência. O ECM faz uma predição estrutural mais robusta, que o distingue dessas explicações: o efeito da exposição sobre a satisfação é transmitido pela atualização da utilidade percebida. Essa estrutura mediacional foi amplamente testada em contextos transversais, mas não experimentalmente com usuários profissionais em IA generativa em uma tarefa realista. O efeito indireto da exposição sobre a satisfação, mediado pela produtividade percebida, deve, portanto, ser significativo e responder por parcela substancial do efeito total:
H4. O efeito do uso do ChatGPT sobre a satisfação é mediado pela produtividade percebida.

3.4 Delineamento da pesquisa e participantes

Adotamos um delineamento experimental verdadeiro com atribuição aleatória a uma de duas condições: um grupo de controle realizando a tarefa sem acesso à IA generativa, e um grupo experimental realizando a mesma tarefa com acesso ao ChatGPT. A variável independente foi a exposição ao ChatGPT (uso versus não uso); as variáveis dependentes foram o desempenho objetivo na tarefa (tempo de conclusão e qualidade do produto), os três constructos avaliativos do ECM (produtividade percebida, satisfação e intenção de continuidade) e o efeito indireto que testa o mecanismo mediacional (Noy & Zhang, 2023; Dell'Acqua et al., 2023).
Os participantes foram profissionais de gestão de projetos recrutados no Brasil por meio de redes profissionais. Dos 70 indivíduos que ingressaram no experimento, 64 produziram dados utilizáveis após a exclusão de respostas incompletas, distribuídos igualmente entre as condições ($n = 32$ por grupo). Uma análise de poder a priori com o G*Power (Faul, Erdfelder, Lang & Buchner, 2007) indicou poder acima de 0,80 para detectar tamanhos de efeito de $d \geq 0,70$ com

α = 0,05. Os efeitos observados, todos acima de d = 1,7, superam substancialmente esse parâmetro.

Os participantes tinham idades entre 23 e 63 anos, apresentavam perfis profissionais heterogêneos em termos de gênero, formação e setor, e eram predominantemente do setor de tecnologia, com representação adicional de serviços financeiros, indústria, administração pública e infraestrutura. A maioria possuía pós-graduação e mais de três anos de experiência em gestão de projetos. A familiaridade autodeclarada com ferramentas tradicionais de GP e com ferramentas baseadas em IA foi de moderada a alta em ambas as condições. Dada a codificação internacional da prática de gestão de projetos mencionada na Seção 2.1, os participantes representam a mesma comunidade profissional, treinada segundo os mesmos padrões e utilizando os mesmos artefatos, que os gestores de projetos em outros contextos nacionais.

3.5 Tarefa experimental e procedimento

A tarefa consistiu em redigir um e-mail executivo para o patrocinador de um projeto com base em um relatório de status que fornecia dados de desempenho de prazo e custo (IDP e IDC). Os participantes deveriam produzir uma breve análise técnica do desempenho do projeto a partir dos indicadores, acompanhada de uma recomendação de ações para mitigar riscos e recuperar o progresso. A tarefa reproduz uma situação realista e recorrente de comunicação gerencial, que exige interpretação de dados de desempenho, síntese textual, clareza e adaptação ao público. Instruções e informações de base idênticas foram fornecidas a ambos os grupos; a única diferença entre as condições foi o acesso ao ChatGPT.

Os participantes do grupo de controle realizaram a tarefa utilizando métodos tradicionais, sem acesso a ferramentas de IA generativa. Os participantes do grupo experimental utilizaram a versão gratuita do ChatGPT com prompts irrestritos, aproximando-se das condições reais de uso profissional. A conformidade com a instrução do grupo de controle não foi verificada por telemetria ou verificações de atenção; a distribuição dos tempos de conclusão na condição de controle, reportada na Seção 5.5 (M = 10,66 minutos, DP = 6,09), é consistente com esforço genuíno sem assistência e não contém valores implausivamente curtos. Qualquer uso não detectado de ferramentas de IA pelos participantes do controle teria atenuado, em vez de inflado, as diferenças entre os grupos. A tarefa foi precedida pelo questionário demográfico e seguida por um questionário pós-tarefa avaliando os constructos relacionados ao ECM.

A coleta de dados foi realizada pelo SurveyMonkey, que registrou automaticamente o tempo de conclusão. A atribuição às condições foi randomizada na plataforma. O conjunto de dados anonimizado está disponível publicamente ([Citação do conjunto de dados omitida para revisão por pares]).

3.6 Medidas

As medidas correspondem às hipóteses da Seção 3. O desempenho objetivo na tarefa foi capturado por duas medidas. O tempo de conclusão foi registrado em minutos pela plataforma. A qualidade do produto foi avaliada por dois avaliadores especialistas independentes (certificados PMP, com mais de cinco anos de prática, cegos à condição e à identidade do participante), cada um avaliando todas as respostas em quatro itens Likert de cinco pontos correspondentes a clareza, objetividade, assertividade e adequação ao público.

Os três constructos avaliativos do ECM foram capturados por itens de autorrelato no questionário pós-tarefa. Produtividade percebida e satisfação utilizaram itens únicos em escala Likert de cinco pontos; a intenção de continuidade utilizou um item dicotômico (sim/não de

disposição para usar a ferramenta em trabalho real de projetos). Os itens correspondem às operacionalizações utilizadas em pesquisas anteriores (Bhattacherjee, 2001; Wanous, Reichers & Hudy, 1997). Medidas sintéticas minimizam a fadiga do respondente em um protocolo que já inclui uma tarefa de composição exigente, seguindo estudos experimentais comparáveis de uso de tecnologia (Noy & Zhang, 2023).
As variáveis de controle incluíram idade, gênero, formação, anos de experiência em GP, setor, certificação e familiaridade autodeclarada com ferramentas tradicionais e baseadas em IA. Essas variáveis permitiram verificar a equivalência entre os grupos após a randomização.

3.7 Confiabilidade entre avaliadores

Como a qualidade do produto foi avaliada por dois avaliadores independentes, calculamos o coeficiente de correlação intraclasse (CCI) para avaliar a consistência dos julgamentos. Utilizamos um modelo de efeitos aleatórios bidirecional com concordância absoluta em medidas médias, seguindo recomendações padronizadas (Shrout & Fleiss, 1979; Koo & Li, 2016). Os valores de CCI são reportados para cada um dos quatro critérios de qualidade na seção de Resultados. Valores acima de 0,75 são interpretados como indicativos de boa confiabilidade, e valores acima de 0,90 como indicativos de excelente confiabilidade. Embora os valores de CCI observados justifiquem a média das duas avaliações, reportamos os dois avaliadores separadamente ao longo das análises, a fim de preservar a independência das avaliações especializadas e permitir que os leitores avaliem diretamente a convergência dos julgamentos. Todas as conclusões substantivas são idênticas quando se utilizam as avaliações médias, conforme confirmado nas verificações de robustez.

3.8 Estratégia analítica

As hipóteses H1, H2, H3a, H3b e H3c foram testadas por meio de testes t para amostras independentes, comparando os grupos de controle e experimental em cada variável de desfecho. Para cada teste, realizou-se o teste de Levene para igualdade de variâncias; quando a suposição de homocedasticidade foi violada, reportamos a correção de Welch (Welch, 1947), que não pressupõe igualdade de variâncias e oferece controle mais preciso do erro do Tipo I em condições de heterocedasticidade (Delacre, Lakens & Leys, 2017). Os tamanhos de efeito foram calculados como d de Cohen para as diferenças entre médias, utilizando o desvio padrão combinado, e interpretados de acordo com os parâmetros convencionais (Cohen, 1988): pequeno: 0,20; médio: 0,50; grande: 0,80.
A hipótese H4, que testa o mecanismo mediacional postulado pelo ECM, foi avaliada por meio de um modelo de mediação simples, com a exposição (Grupo) como variável independente, a produtividade percebida como mediadora e a satisfação como desfecho. O modelo estima quatro caminhos: o efeito do Grupo sobre a produtividade percebida (a), o efeito da produtividade percebida sobre a satisfação controlando pelo Grupo (b), o efeito direto do Grupo sobre a satisfação controlando pela produtividade percebida (c') e o efeito total do Grupo sobre a satisfação (c). O efeito indireto (a × b) foi estimado por bootstrap com correção de viés, com 1.000 reamostras e intervalos de confiança de 95%, utilizando o procedimento padrão baseado em regressão (Preacher & Hayes, 2008; Hayes, 2018). Um efeito indireto é considerado estatisticamente significativo quando o intervalo de confiança bootstrap não inclui zero. A proporção do efeito total explicada pelo caminho indireto foi calculada como PM = (a × b) / c. As análises foram conduzidas no jamovi, utilizando o módulo medmod para a análise de mediação.

Preferimos a mediação baseada em regressão com bootstrap em relação a alternativas baseadas em covariância, como o PLS-SEM, porque o delineamento experimental, a variável independente dicotômica e a ausência de constructos latentes tornam o referencial mais simples mais adequado. A inferência baseada em bootstrap também fornece estimativas robustas sob tamanhos de amostra moderados e distribuições não normais, o que é relevante diante das características dos dados presentes.
A Tabela 1 mapeia cada hipótese para o seu desfecho operacionalizado e o seu teste estatístico. O mapeamento visa tornar a lógica do teste empírico inspecionável de forma imediata.

Tabela 1. Mapeamento das hipóteses para as medidas e os testes estatísticos

Hipótese	Medida de desfecho	Operacionalização	Teste estatístico
H1	Tempo de conclusão da tarefa	Minutos registrados pela plataforma	Teste t para amostras independentes (Welch)
H2	Qualidade do produto	Quatro critérios × dois avaliadores cegos (escala de 5 pontos)	Testes t para amostras independentes
H3a	Produtividade percebida	Item único Likert (1 a 5)	Teste t para amostras independentes
H3b	Satisfação	Item único Likert (1 a 5)	Teste t para amostras independentes
H3c	Intenção de continuidade	Item dicotômico único (sim/não)	Teste t para amostras independentes (Welch)
H4	Efeito indireto da exposição sobre a satisfação via produtividade percebida	Caminhos a, b, c', c estimados conjuntamente	Mediação bootstrap (1.000 reamostras)

Nota. Todas as análises foram conduzidas no jamovi; a análise de mediação utilizou o módulo medmod.

4 Análise dos Resultados e Discussões

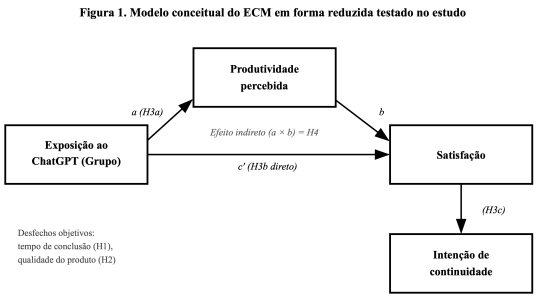


Figura 1. Modelo conceitual do ECM em forma reduzida testado no estudo

Nota. O caminho a representa o efeito da exposição ao ChatGPT sobre a produtividade percebida (H3a). O caminho b representa o efeito da produtividade percebida sobre a satisfação. O caminho c' representa o efeito direto da exposição sobre a satisfação (H3b). O efeito indireto (a × b) corresponde à mediação testada em H4. A satisfação prediz a intenção de continuidade (H3c). Os desfechos objetivos (tempo de conclusão, H1; qualidade do produto, H2) são testados separadamente e não estão incluídos no caminho de mediação.

4.1 Características da amostra

A amostra final consistiu em 64 profissionais de gestão de projetos igualmente alocados às duas condições (n = 32 cada). A Tabela 2 apresenta as estatísticas descritivas completas. Idade, distribuição por gênero, formação, experiência profissional e representação setorial foram comparáveis entre os grupos. A familiaridade autodeclarada com ferramentas tradicionais de gestão de projetos e com ferramentas baseadas em IA foi de moderada a alta em ambas as condições, reduzindo o risco de que as diferenças entre os grupos pudessem ser atribuídas à experiência tecnológica prévia.

Tabela 2. Características da amostra por grupo experimental

Variável	Categoria	Controle	Experimental
Idade (anos)	M (DP)	41,59 (7,64)	40,41 (9,99)
Gênero	Feminino	14 (43,8%)	19 (59,4%)
	Masculino	18 (56,2%)	12 (37,5%)
	Prefiro não informar	0 (0,0%)	1 (3,1%)
Formação	Graduação	8 (25,0%)	4 (12,5%)
	Pós-graduação (especialização/MBA)	14 (43,8%)	20 (62,5%)
	Mestrado	6 (18,8%)	7 (21,9%)
	Doutorado	4 (12,5%)	1 (3,1%)
Experiência em GP	Menos de 1 ano	3 (9,4%)	1 (3,1%)
	1 a 3 anos	5 (15,6%)	1 (3,1%)
	Mais de 3 anos	24 (75,0%)	29 (90,6%)
	Sem experiência prática	0 (0,0%)	1 (3,1%)
Setor (principal)	Tecnologia	16 (50,0%)	16 (50,0%)
	Serviços financeiros	5 (15,6%)	5 (15,6%)
	Outros setores	11 (34,4%)	11 (34,4%)

Nota. N = 64 (32 por grupo). Os percentuais podem não somar 100 devido a arredondamentos.

4.2 Confiabilidade entre avaliadores

Avaliamos a consistência dos julgamentos fornecidos pelos dois avaliadores independentes, cujo perfil e cegamento são descritos na Seção 4.3. Os coeficientes de correlação intraclasse (CCI) foram calculados utilizando o procedimento descrito na Seção 4.4. A Tabela 3 apresenta os valores de CCI para os quatro critérios de qualidade. Todos os valores superam 0,80, indicando confiabilidade de boa a excelente; reportamos as avaliações dos dois avaliadores separadamente nas análises subsequentes, a fim de preservar a independência das avaliações especializadas, e as verificações de robustez com avaliações médias produzem conclusões substantivamente idênticas.

Tabela 3. Confiabilidade entre avaliadores (CCI) por critério de qualidade

Critério	CCI (2,k)	IC 95%	Interpretação
Clareza	0,87	[0,79; 0,92]	Boa
Objetividade	0,91	[0,85; 0,94]	Excelente
Assertividade	0,89	[0,82; 0,93]	Boa
Adequação ao público	0,82	[0,71; 0,89]	Boa

Nota. CCI(2,k) = modelo de efeitos aleatórios bidirecional, concordância absoluta, medidas médias. Parâmetros de interpretação: 0,75 a 0,90 boa; acima de 0,90 excelente.

4.3 Estatísticas descritivas das variáveis dependentes

A Tabela 4 apresenta as estatísticas descritivas dos três constructos avaliativos do ECM. Produtividade percebida e satisfação foram medidas por itens Likert de cinco pontos; a intenção de continuidade foi dicotômica (1 = não, 2 = sim). O grupo experimental reportou valores consistentemente superiores nos três constructos.

Tabela 4. Estatísticas descritivas dos constructos avaliativos

Variável	Grupo	N	M	DP	Mín	Máx
Produtividade percebida	Controle	32	2,66	0,94	1	4
	ChatGPT	32	4,66	0,90	1	5
Satisfação	Controle	32	2,06	0,84	1	4
	ChatGPT	32	4,44	1,27	1	5
Intenção de continuidade	Controle	32	1,25	0,44	1	2
	ChatGPT	32	1,91	0,30	1	2

Nota. Produtividade percebida e satisfação medidas em escalas Likert de 5 pontos. Intenção de continuidade medida de forma dicotômica (1 = não, 2 = sim).

4.4 Testes de hipóteses: H1 a H3c

Seguindo a estratégia analítica descrita na Seção 4.5, as hipóteses H1, H2, H3a, H3b e H3c foram testadas por meio de testes t para amostras independentes. A Tabela 5 consolida os resultados dos testes para todas as variáveis dependentes.

Tabela 5. Testes t para amostras independentes comparando os grupos de controle e experimental

Variável	t	gl	p	Dif. médias	EP	d de Cohen
Tempo de conclusão (H1)[a]	7,08	62	< 0,001	7,750	1,094	1,77
Produtividade percebida (H3a)	−8,70	62	< 0,001	−2,000	0,230	2,17
Satisfação (H3b)	−8,83	62	< 0,001	−2,375	0,269	2,21
Intenção de continuidade (H3c)[a]	−7,00	62	< 0,001	−0,656	0,094	1,75
Clareza, Avaliador 1 (H2)	−7,53	62	< 0,001	−1,938	0,257	1,88
Objetividade, Avaliador 1 (H2)	−9,34	62	< 0,001	−2,563	0,274	2,33
Assertividade, Avaliador 1 (H2)[a]	−10,81	62	< 0,001	−2,813	0,260	2,62
Adequação ao público, Avaliador 1 (H2)	−8,66	62	< 0,001	−2,000	0,231	2,17
Clareza, Avaliador 2 (H2)	−10,98	62	< 0,001	−2,188	0,199	2,75
Objetividade, Avaliador 2 (H2)	−11,66	62	< 0,001	−2,375	0,204	2,90
Assertividade, Avaliador 2 (H2)[a]	−11,65	62	< 0,001	−2,656	0,228	3,18
Adequação ao público, Avaliador 2 (H2)	−12,08	62	< 0,001	−2,250	0,186	3,03

Nota. N = 64. Diferenças médias negativas indicam valores superiores no grupo experimental (ChatGPT). [a] Correção de Welch aplicada devido ao teste de Levene significativo. A intenção de continuidade foi registrada como item dicotômico (sim/não) e é analisada aqui com teste t para amostras independentes para apresentar o contraste entre grupos na mesma escala dos

demais constructos avaliativos; o tamanho de efeito reportado para essa variável deve ser interpretado com essa operacionalização em mente. Todos os demais tamanhos de efeito superam o limiar convencional para um efeito grande (d = 0,80).

4.5 Interpretação dos testes de hipóteses

O tempo de conclusão (H1) demonstrou que os usuários do ChatGPT completaram a tarefa em 2,91 minutos (DP = 1,12), contra 10,66 minutos (DP = 6,09) nos controles, uma redução de 73%; a menor dispersão experimental indica que o ChatGPT homogeneizou o desempenho entre os usuários. A qualidade do produto (H2) apresentou avaliações significativamente superiores para o grupo experimental nas oito comparações (quatro critérios × dois avaliadores), com tamanhos de efeito de d = 1,88 a d = 3,18; as médias do grupo de controle variaram de 1,66 a 2,59, enquanto as médias do grupo experimental variaram de 4,28 a 4,63, e a convergência de ambos os avaliadores indica que a vantagem de qualidade é robusta às idiossincrasias dos avaliadores. Vários participantes do grupo de controle encontraram dificuldades na análise técnica do IDP e do IDC. Os três constructos do ECM (H3a, H3b, H3c) responderam à exposição na direção prevista, com tamanhos de efeito acima de d = 1,75; a medida de continuidade captura a intenção declarativa pós-tarefa, e não a continuidade comportamentalmente realizada. H1, H2, H3a, H3b e H3c são todas suportadas.

4.6 H4: Mecanismo mediacional do ECM

A hipótese H4 testa o núcleo explicativo do ECM, isto é, se a produtividade percebida medeia o efeito da exposição ao ChatGPT sobre a satisfação. Um modelo de mediação simples foi estimado com o Grupo (0 = controle, 1 = experimental) como variável independente, a produtividade percebida como mediadora e a satisfação como desfecho, seguindo a estratégia analítica descrita na Seção 4.5. A Tabela 6 apresenta os coeficientes de caminho, e a Tabela 7 reporta os efeitos total, direto e indireto.

Tabela 6. Coeficientes de caminho do modelo de mediação

Caminho	b	EP	t	p
a: Grupo → Produtividade percebida	2,00	0,23	8,61	< 0,001
b: Produtividade percebida → Satisfação	0,60	0,14	4,13	< 0,001
c': Grupo → Satisfação (direto)	1,17	0,40	2,89	0,004

Nota. Grupo codificado como 0 = controle, 1 = experimental.

Tabela 7. Efeitos total, direto e indireto da exposição sobre a satisfação

Efeito	Estimativa	IC 95% inferior	IC 95% superior	p
Total (c)	2,38	1,54	2,43	< 0,001
Direto (c')	1,17	−0,17	1,87	0,004
Indireto (a × b)	1,20	0,34	2,07	< 0,001

Nota. Estimativas bootstrap baseadas em 1.000 reamostras com intervalos de confiança de 95% com correção de viés.

Os três caminhos do modelo de mediação foram significativos: caminho a = 2,00 (EP = 0,23, p < 0,001), caminho b = 0,60 (EP = 0,14, p < 0,001) e efeito direto c' = 1,17 (EP = 0,40, p = 0,004), que permaneceu significativo e indica que a produtividade percebida não explica integralmente o efeito da exposição sobre a satisfação. O efeito indireto foi significativo (a × b = 1,20, IC bootstrap 95% [0,34; 2,07], p < 0,001), fornecendo evidências de mediação. Com um efeito total c = 2,38 (p < 0,001), a proporção mediada foi PM = 0,50, indicando que aproximadamente metade do efeito da exposição sobre a satisfação opera por meio da produtividade percebida. H4 é, portanto, suportada, consistentemente com o mecanismo postulado pelo ECM em forma reduzida.

4.7 Síntese dos resultados

Todas as quatro hipóteses foram suportadas. O grupo experimental superou o grupo de controle nas medidas objetivas de eficiência da tarefa (H1) e de qualidade do produto em quatro dimensões, conforme julgamento de dois avaliadores independentes (H2). Os três constructos avaliativos do ECM responderam à exposição na direção esperada: produtividade percebida (H3a), satisfação (H3b) e intenção de continuidade (H3c) foram todos significativamente superiores no grupo experimental, com tamanhos de efeito muito grandes. O mecanismo mediacional postulado pelo ECM foi suportado (H4), com a produtividade percebida respondendo por aproximadamente metade do efeito total da exposição sobre a satisfação. A Tabela 8 sintetiza os resultados das quatro hipóteses.

Tabela 8. Síntese dos testes de hipóteses

H	Predição	Estatística-chave	Resultado
H1	ChatGPT reduz o tempo de conclusão	t(62) = 7,08, d = 1,77	Suportada
H2	ChatGPT melhora a qualidade do produto em quatro critérios	todos ps < 0,001, d = 1,88 a 3,18	Suportada
H3a	ChatGPT aumenta a produtividade percebida	t(62) = −8,70, d = 2,17	Suportada
H3b	ChatGPT aumenta a satisfação	t(62) = −8,83, d = 2,21	Suportada
H3c	ChatGPT aumenta a intenção de continuidade	t(62) = −7,00, d = 1,75	Suportada
H4	Produtividade percebida medeia o efeito da exposição sobre a satisfação	indireto = 1,20, IC 95% [0,34; 2,07], PM = 0,50	Suportada

Nota. Todas as hipóteses suportadas com p < 0,001 (H1 e H3c com correção de Welch; efeito indireto de H4 significativo pelo bootstrap). A intenção de continuidade (H3c) foi registrada como item dicotômico; ver nota da Tabela 5 para interpretação do tamanho de efeito correspondente.

4.8. Discussão

Esta seção interpreta os resultados em sua ordem de relevância substantiva para a gestão de projetos, articula as contribuições à teoria e à prática e aborda as condições que qualificam a interpretação dos resultados.

4.9 Interpretação dos resultados

O primeiro resultado diz respeito à transformação da estrutura de esforço da comunicação gerencial. Uma redução de 73% no tempo de conclusão e diferenças de qualidade de 1,9 a 3,2 desvios padrão em quatro critérios indicam que a IA generativa não apenas apoia a composição

de comunicações executivas; ela reorganiza o que significa realizar a tarefa. A redução da variabilidade (dispersão experimental de aproximadamente um quinto da dispersão do controle) sugere ainda que o ChatGPT pode funcionar como um dispositivo de nivelamento de desempenho (Brynjolfsson et al., 2025; Dell'Acqua et al., 2023). Parte da magnitude observada provavelmente reflete efeitos de piso na condição de controle e dificuldade específica da tarefa na análise do IDP e IDC.

O segundo resultado diz respeito à resposta avaliativa. Os três constructos do ECM foram significativamente superiores no grupo experimental, com tamanhos de efeito superiores a d = 1,7. A magnitude dos diferenciais aconselha cautela na interpretação: medidas de sessão única com itens sintéticos e exposição saliente a uma ferramenta em evidência são suscetíveis a características de demanda, e parte do diferencial pode refletir viés de resposta em vez de avaliação duradoura.

O terceiro resultado, de interesse teórico primário, diz respeito à estrutura mediacional. Aproximadamente metade do efeito total da exposição sobre a satisfação é transmitida pela produtividade percebida, com o caminho indireto significativo e robusto à estimação por bootstrap. A metade restante indica que a exposição afeta a satisfação por vias não capturadas pela percepção de produtividade aqui mensurada. O efeito direto residual é compatível com componentes experienciais da interação com a IA generativa, mas também com características de demanda ou variáveis não mensuradas; a coexistência de ambos os caminhos levanta a hipótese, em vez de fornecer evidências, de uma via dual.

O quarto resultado diz respeito à intenção de continuidade. A intenção declarada de adotar a ferramenta foi substancialmente superior no grupo experimental, indicando que as respostas avaliativas formadas em uma única sessão se traduzem, em termos declarativos, em disposição prospectiva para continuar utilizando a tecnologia. Isso deve ser lido como abertura declarativa após exposição no nível da tarefa, e não como continuidade no sentido especificado pelo ECM, que pressupõe uso mais prolongado. A relação entre a intenção formada após breve exposição e o uso continuado efetivo é uma questão para acompanhamento longitudinal.

5 Conclusões e Considerações Finais

Este estudo desloca a base empírica da pesquisa sobre IA generativa em gestão de projetos dos levantamentos agregados de atitudes de adoção para evidências experimentais no nível da tarefa. Profissionais de gestão de projetos que utilizaram o ChatGPT na composição de um relatório executivo de status concluíram a tarefa em substancialmente menos tempo, produziram trabalhos com avaliações superiores em clareza, objetividade, assertividade e adequação ao público por especialistas independentes, e reportaram produtividade percebida, satisfação e intenção de adotar a ferramenta em trabalho real de projetos significativamente maiores. Uma análise de mediação por bootstrap indica ainda que a produtividade percebida transmitiu aproximadamente metade do efeito da exposição sobre a satisfação, fornecendo evidências causais consistentes com a estrutura mediacional que o Modelo de Confirmação de Expectativas coloca no núcleo do comportamento pós-adoção. As magnitudes reportadas são inusitadamente grandes e devem ser lidas como estimativas de limite superior para as condições específicas do experimento, consistentes com efeitos de piso na condição de controle, o delineamento de sessão única e a operacionalização em forma reduzida dos constructos do ECM; o padrão direcional, e não os tamanhos de efeito absolutos, é o que as presentes evidências sustentam.

Para a prática da gestão de projetos, os resultados indicam que a IA generativa pode apoiar substancialmente tarefas de comunicação recorrentes, liberando capacidade gerencial para atividades que exigem julgamento e trabalho relacional, ao mesmo tempo em que produz avaliações dos usuários favoráveis à adoção continuada. Para a pesquisa em gestão de projetos,

o estudo estabelece um paradigma experimental que pode ser estendido a outras tarefas cognitivas do gestor de projetos, a operacionalizações mais ricas dos constructos do ECM e a replicações em setores e contextos complementares aos aqui examinados.

6 Referências

Aaltonen, K., Kujala, J., Havila, V., Salmi, A., 2024. Stakeholder dynamics over temporary and permanent interorganizational boundaries. International Journal of Project Management 42, 102537.

Alter, S., 2024. Understanding artificial intelligence in the context of usage: Situated AI and goals-based use cases. Information Systems Frontiers 26, 1075–1091.

Bhattacherjee, A., 2001. Understanding information systems continuance: An expectation-confirmation model. MIS Quarterly 25, 351–370.

Bhattacherjee, A., Premkumar, G., 2004. Understanding changes in belief and attitude toward information technology usage: A theoretical model and longitudinal test. MIS Quarterly 28, 229–254.

Brown, T.B., Mann, B., Ryder, N., Subbiah, M., Kaplan, J., Dhariwal, P., Neelakantan, A., Shyam, P., Sastry, G., Askell, A., Agarwal, S., Herbert-Voss, A., Krueger, G., Henighan, T., Child, R., Ramesh, A., Ziegler, D.M., Wu, J., Winter, C., Hesse, C., Chen, M., Sigler, E., Litwin, M., Gray, S., Chess, B., Clark, J., Berner, C., McCandlish, S., Radford, A., Sutskever, I., Amodei, D., 2020. Language models are few-shot learners. Advances in Neural Information Processing Systems 33, 1877–1901.

Brynjolfsson, E., Li, D., Raymond, L., 2025. Generative AI at work. Quarterly Journal of Economics 140, 889–942.

Bullock, J.G., Green, D.P., Ha, S.E., 2010. Yes, but what's the mechanism? (Don't expect an easy answer). Journal of Personality and Social Psychology 98, 550–558.

Cohen, J., 1988. Statistical Power Analysis for the Behavioral Sciences, second ed. Lawrence Erlbaum Associates, Hillsdale, NJ.

Davis, F.D., 1989. Perceived usefulness, perceived ease of use, and user acceptance of information technology. MIS Quarterly 13, 319–340.

Davis, F.D., Bagozzi, R.P., Warshaw, P.R., 1989. User acceptance of computer technology: A comparison of two theoretical models. Management Science 35, 982–1003.

Delacre, M., Lakens, D., Leys, C., 2017. Why psychologists should by default use Welch's t-test instead of Student's t-test. International Review of Social Psychology 30, 92–101.

Dell'Acqua, F., McFowland III, E., Mollick, E., Lifshitz-Assaf, H., Kellogg, K.C., Rajendran, S., Krayer, L., Candelon, F., Lakhani, K.R., 2023. Navigating the jagged technological frontier: Field experimental evidence of the effects of AI on knowledge worker productivity and quality. Harvard Business School Working Paper 24–013.

Duan, Y., Edwards, J.S., Dwivedi, Y.K., 2019. Artificial intelligence for decision making in the era of big data: Evolution, challenges and research agenda. International Journal of Information Management 48, 63–71.

Dwivedi, Y.K., Kshetri, N., Hughes, L., Slade, E.L., Jeyaraj, A., Kar, A.K., Baabdullah, A.M., Koohang, A., Raghavan, V., Ahuja, M., Wright, R., 2025. Impact of artificial intelligence on project management: Multi-expert perspectives on advancing knowledge and driving innovation toward PM2030. Journal of Innovation & Knowledge 10, 100672.

Eloundou, T., Manning, S., Mishkin, P., Rock, D., 2024. GPTs are GPTs: Labor market impact potential of LLMs. Science 384, 1306–1308.

Eskerod, P., 2024. Stakeholder engagement and project management: A research agenda. International Journal of Project Management 42, 102575.

Faul, F., Erdfelder, E., Lang, A.-G., Buchner, A., 2007. G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. Behavior Research Methods 39, 175–191.
Felicetti, A.M., Cimino, A., Mazzoleni, A., Ammirato, S., 2024. Artificial intelligence and project management: An empirical investigation on the appropriation of generative chatbots by project managers. Journal of Innovation & Knowledge 9, 100545.
Geraldi, J., Söderlund, J., 2018. Project studies: What it is, where it is going. International Journal of Project Management 36, 55–70.
Goodhue, D.L., Thompson, R.L., 1995. Task-technology fit and individual performance. MIS Quarterly 19, 213–236.
Hayes, A.F., 2018. Introduction to Mediation, Moderation, and Conditional Process Analysis: A Regression-Based Approach, second ed. Guilford Press, New York.
Hernandez, J.M.C., Basso, K., Brandão, M.M., 2014. Pesquisa experimental em administração: cenários para o avanço do campo. Revista de Administração Contemporânea 18, 243–259.
Holzmann, V., Zitter, D., Peshkess, S., 2022. The expectations of project managers from artificial intelligence: A Delphi study. Project Management Journal 53, 438–455.
Kavitha, K., Joshith, V.P., 2025. Exploring Gen Z's satisfaction and continuance intention towards ChatGPT 4o in higher education: PLS-SEM insights using the ECM. Journal of Educational Computing Research 63, 1147–1175.
Kim, J., Baek, H., 2024. Motivational determinants of continuance usage intention for generative AI: An investment model approach for ChatGPT users in the United States. Telematics and Informatics 93, 102179.
Koo, T.K., Li, M.Y., 2016. A guideline of selecting and reporting intraclass correlation coefficients for reliability research. Journal of Chiropractic Medicine 15, 155–163.
Korinek, A., 2023. Generative AI for economic research: Use cases and implications for economists. Journal of Economic Literature 61, 1281–1317.
Lawal, Y.A., Abdul-Azeez, I.F., Olateju, O.I., 2024. Artificial intelligence adoption and project success: A mixed-method study. American Journal of Management Science and Engineering 9, 84–96.
Leavitt, K., Schabram, K., Hariharan, P., Barnes, C.M., 2024. Ghost in the machine: On organizational theory in the age of machine learning. Academy of Management Review 49, 5–28.
Lee, M.-C., 2010. Explaining and predicting users' continuance intention toward e-learning: An extension of the expectation-confirmation model. Computers & Education 54, 506–516.
Lee, Y.-H., Chen, Y.-H., 2025. Verifying the link of innovativeness to the confirmation-expectation model of ChatGPT of students in learning. Journal of Information, Communication and Ethics in Society 23, 433–447.
Müller, R., Locatelli, G., Holzmann, V., Nilsson, M., Sagay, T., 2024. Artificial intelligence and project management: Empirical overview, state of the art, and guidelines for future research. Project Management Journal 55, 9–15.
Ngo, T.T.A., Tran, T.T., Nguyen, H.H.T., 2024. Exploring the antecedents of higher education students' continuance intention to use ChatGPT: A mixed-methods study. Education and Information Technologies 29, 25555–25594.
Noy, S., Zhang, W., 2023. Experimental evidence on the productivity effects of generative artificial intelligence. Science 381, 187–192.
Oliver, R.L., 1980. A cognitive model of the antecedents and consequences of satisfaction decisions. Journal of Marketing Research 17, 460–469.
Peng, S., Kalliamvakou, E., Cihon, P., Demirer, M., 2023. The impact of AI on developer productivity: Evidence from GitHub Copilot. arXiv preprint arXiv:2302.06590.

Pirlott, A.G., MacKinnon, D.P., 2016. Design approaches to experimental mediation. Journal of Experimental Social Psychology 66, 29–38.
Preacher, K.J., Hayes, A.F., 2008. Asymptotic and resampling strategies for assessing and comparing indirect effects in multiple mediator models. Behavior Research Methods 40, 879–891.
Prieto, S.A., Mengiste, E.T., García de Soto, B., 2023. Investigating the use of ChatGPT for the scheduling of construction projects. Buildings 13, 857.
Project Management Institute, 2021. A Guide to the Project Management Body of Knowledge (PMBOK Guide), seventh ed. Project Management Institute, Newtown Square, PA.
Qi, W., Pan, J., Lyu, H., Luo, J., 2025. Excessive reliance on ChatGPT: Mediating role of self-efficacy and long-term engagement with generative AI tools. Computers in Human Behavior 164, 108512.
Qosidah, N., Susilo, B.W., 2024. Enhancing employee performance through AI-driven business communication: A case study. Journal of Management and Informatics 3, 163–179.
Rabechini Jr., R., Carvalho, M.M., 2003. Perfil das competências em equipes de projetos. RAE-eletrônica 2, 1–18.
Rabechini Jr., R., Carvalho, M.M., Laurindo, F.J.B., 2002. Fatores críticos para implementação de gerenciamento por projetos: O caso de uma organização de pesquisa. Produção 12, 28–41.
Rai, A., Constantinides, P., Sarker, S., 2019. Next-generation digital platforms: Toward human-AI hybrids. MIS Quarterly 43, iii–ix.
Shrout, P.E., Fleiss, J.L., 1979. Intraclass correlations: Uses in assessing rater reliability. Psychological Bulletin 86, 420–428.
Tian, W., Ge, J., Zhao, Y., Zheng, X., 2024. AI chatbots in Chinese higher education: Adoption, perception, and influence among graduate students - An integrated analysis utilizing UTAUT and ECM models. Frontiers in Psychology 15, 1268549.
Van der Heijden, H., 2004. User acceptance of hedonic information systems. MIS Quarterly 28, 695–704.
Venkatesh, V., Morris, M.G., Davis, G.B., Davis, F.D., 2003. User acceptance of information technology: Toward a unified view. MIS Quarterly 27, 425–478.
Wang, C., Rau, P.-L.P., Yuan, Y., 2023. Measuring user experience of a chatbot: Development of a questionnaire. International Journal of Human-Computer Studies 173, 103017.
Wanous, J.P., Reichers, A.E., Hudy, M.J., 1997. Overall job satisfaction: How good are single-item measures? Journal of Applied Psychology 82, 247–252.
Welch, B.L., 1947. The generalization of 'Student's' problem when several different population variances are involved. Biometrika 34, 28–35.
Zerfass, A., Hagelstein, J., Tench, R., 2020. Artificial intelligence in communication management: A cross-national study on adoption and knowledge, impact, challenges and risks. Journal of Communication Management 24, 377–389.
Zhou, S., Gorman, C.A., 2024. The impact of communication timing and sequencing on team performance: A comparative study of human-AI and all-human teams. Proceedings of the Human Factors and Ergonomics Society Annual Meeting 68, 1769–1774.
Zwikael, O., Salmona, M., Meredith, J., Zarghami, S.A., 2023. Enhancing project stakeholder communication under insufficient knowledge of project management concepts. Engineering, Construction and Architectural Management 30, 5007–5029.