

Research Article

Beyond Self-Report: Preliminary AI Analysis of Language Related to the Burnout/Depression Relationship

Scott T. Meier¹, Susan Davis²

1. University at Buffalo, United States; 2. Independent researcher

Researchers have documented moderate to large correlations between scores on burnout and depression measures, fueling debate over whether the two constructs are distinct. A key limitation of this literature is its near-exclusive reliance on self-report measures, which are vulnerable to shared method variance, mono-method bias, and intracategory variability in item interpretation. This exploratory study examines whether burnout and depression can be distinguished at the level of natural language using Large Language Models (LLMs). Drawing on publicly available qualitative data, transcripts and verbatim excerpts were analyzed from frontline healthcare professionals (11 participants, 1,017 speech turns). Using detailed, theory-based prompts, two LLMs (ChatGPT and ClaudeAI) independently classified participant statements as reflecting burnout, depression, both constructs, or neither. Healthcare workers' speech primarily reflected burnout, although a subset of individuals also evidenced depressive language. Substantial individual differences emerged in the degree of overlap between burnout and depression, suggesting that this relationship may be idiographic rather than uniform across individuals. A dose-response interpretation may explain the relative absence of depression among this group of less-experienced healthcare workers. LLM analysis appears to offer a complementary approach that may help clarify the dynamic relationship between burnout and depression while addressing limitations inherent in self-reports.

Correspondence: papers@team.qeios.com — Qeios will forward to the authors

Introduction

Over the past 40 years, burnout researchers have sought to interpret the substantial overlap observed between scores on burnout and depression self-report scales ^{[1][2]}. Moderate to large correlations between burnout and depression have been consistently observed between these measures ^{[3][4][5][6]}. A review of 92 studies published between 1984 and 2014 found a comparable degree of burnout–depression overlap across both cross-sectional and longitudinal designs ^[3]. Similarly, a meta-analysis of 67 studies published between 2007 and 2018 found an average correlation of .52 between burnout and depression ^[4].

Despite similarities of findings, researchers have reached different conclusions regarding the relationship between burnout and depression. One argument is that burnout and depression reflect a single underlying construct ^{[3][6]}. In line with this perspective, the 9-item Occupational Depression Inventory (ODI) ^{[6][7]} was created to assess depressive symptoms that individuals attribute to their occupational context. Another new measure of burnout was recently developed to include depressive symptom items ^[8]. Post item-level analyses, test developers concluded that "burnout is unlikely to be synonymous with depression but...a burnout syndrome is commonly if not always accompanied by some depressive symptoms" (p. 6) ^{[8][9]}. Finally, Koutsimani et al. examined correlations among burnout, depression, and anxiety measures and concluded that burnout and depression are more likely to be two different constructs rather than one, with no conclusive overlap among the three ^[4].

Less attention has been paid in the literature to methodological factors that might influence the observed burnout–depression correlation. Evidence exists for mono-operations bias since the Maslach Burnout Inventory (MBI) has been employed in approximately 80% of studies ^[3]. Since burnout and depression are largely measured via self-report, research on burnout also suffers from mono-method bias ^[10]. Shared method variance from self-reports, the degree to which the common measurement methodology contributes to test scores, likely inflates the burnout/depression correlation. In addition, research suggests that individuals often interpret items on self-report psychological tests with different meanings than test developers intended ^[11]. This problem has been labeled intracategory variability, indicating that test-takers interpret and respond to self-report items based on widely varying personal experiences ^[12]. When asked to report a recent "serious illness," for example, respondents will describe events ranging from routine flu to a heart attack. The criteria individuals use for answering health-related items on self-report measures can span from the catastrophic to the trivial ^[12].

Identification of Negative Affect Via Linguistic Analyses

Researchers employ AI and other automated methods to identify psychological constructs and mental health data in text ^{[13][14][15]}. Shapira et al. (2021) examined linguistic features reflecting psychotherapy clients' distress using computerized text analysis ^[14]. Using data from 58 clients and 729 transcribed sessions, they found that fewer first-person singular words and fewer negative emotion words were associated with lower distress in both the same and next session. Decreases in the use of first-person singular words from early to later stages of therapy were related to improved treatment outcomes.

Large language models (LLMs) have been employed to investigate psychological constructs that potentially contribute to ceiling effects observed in working alliance measures ^[16]. Using a classic psychotherapy transcript (the Rogers–Gloria session) ^[17], they applied a psychometric framework to evaluate whether LLMs could detect four theoretically relevant psychological constructs potentially linked to early alliance formation (i.e., client distress, social anxiety, attachment style, and positive transference). Using detailed, construct-specific prompts, ChatGPT analyzed 72 client speech turns; the analysis was then repeated three weeks later to test temporal stability and cross-validated using a different LLM (Claude AI). The AI identified numerous statements consistent with psychological distress and social anxiety, with these constructs showing the highest rates of unique classification, replication, and cross-model agreement ^[16].

When evaluating a large quantity of text, an AI analysis has the potential for greater consistency, cost-effectiveness, and potentially less bias than a human rater ^[18]. Linguistic analysis using AI is also unobtrusive, avoids reactivity effects, and can be performed without human raters ^[19]. AI language analyses can also provide a person-centered idiographic approach ^[20] that assesses an individual's unique characteristics over time that may be unrelated to nomothetic traits.

Methods

A university librarian with expertise in online databases and ChatGPT were consulted to identify studies of persons who describe their experiences of depression and burnout in publicly available qualitative data. A major obstacle to obtaining complete transcripts is that most qualitative studies do not make their transcripts publicly available because of confidentiality promised to research participants. Relative to the purposes of this research, only three studies were found ^{[21][22][23]}. Fullen ^[21] studied individuals diagnosed with clinical depression who reported their perceptions of depressive symptoms, social media

use, and interpersonal relations. Because only 49 extracted quotes from interviews were available, this sample was deemed too small for language analysis of burnout and depression. Similarly, Li et al. [22] conducted semi-structured interviews with 14 operating room (OR) nurses who worked in a large general Chinese hospital. Only 43 statements were available in the study.

Smith et al. [23] employed semi-structured interviews with health and social care professionals to explore barriers and facilitators regarding expressive writing about occupational stress and well-being. Eleven UK-based frontline health and social care workers from psychological and mental health were recruited through professional networks and social media. Participants were 82% female with a mean age of 31.55 years; most participants indicated they had only a few years of work experience. Containing a total of 1,017 speech turns, the interviews were posted online at <https://osf.io/zfmtk/files/guqnv>; permission was obtained to employ this language sample for analysis. For each interview, the transcript designates each speech turn as uttered by either the interviewer or the healthcare worker. Each speech turn consists of one or multiple statements that can include sentences as well as phrases that potentially could be classified as belonging to more than one construct.

Construct Analysis Prompt and AI Applications

Just as the development of psychological tests rests on initial specification of constructs that guide item and task construction [24], specific wording of the prompts employed to query LLMs strongly influences subsequent output. Consequently, considerable effort was devoted to writing detailed prompt instructions for use with study transcripts. Existing research literature provides indications about how burnout and depression might be distinguished in terms of expressed negative affect. Differences may be apparent in emotional tone, focus of attention, cognitive style, and social language [3][4][25][26][27][28]. Burned out individuals, for example, may be more likely to employ words like *exhausted*, *frustration*, and *cynical*, specific to their work context. In contrast, depressed persons' language may include *hopeless*, *worthless*, and *pointless*.

Research findings guided the creation of a detailed prompt (see Appendix A) employed for LLM analysis of transcripts. The AI was instructed to classify each separate piece of text as "reflecting burnout, depression, both burnout and depression, or neither." The prompt then provided multiple examples of burnout versus depression language for the AI program to employ for its analysis.

Participant speech turns consisted of individual segments as designated by the respective researchers. Examples include "I always want to do everything perfectly, so I constantly feel that what I do is never

good enough" and "The hardest part is not the regular work but completing all those additional tasks. At first, it gave me a sense of accomplishment, but over time, the imbalance between effort and reward made me feel resentful."

Large Language Models

An advanced Large Language Model (LLM) based on a knowledge base developed across multiple domains, the ChatGPT program was designed for natural language understanding and generation. ChatGPT has been employed for multiple purposes, including content creation, code generation, and qualitative analysis [29]. OpenAI has introduced a newer model, GPT-4o, which began replacing GPT-4 in ChatGPT as of April 30, 2025. Model sub-versions, temperature parameters, and token-handling settings reflected standard system settings. Similarly, Claude AI is an advanced Large Language Model with training across diverse domains and disciplines. As Claude 4 Opus/Sonnet 4.1, its training data extends through January 2025. Anthropic (claude.ai) characterizes Claude as offering nuanced understanding of complex topics across academic, professional, and creative domains.

Preprocessing procedures involved (a) removing all interviewer content from the transcripts, leaving the LLMs to analyze only participant speech turns, and (b) labelling each participant speech turn with a consecutive number for identification purposes. For each study, the LLMs analyzed the participant transcripts during a single run.

Research Questions

This study's sample of professionals employed in health and social care may potentially evidence both depression and burnout constructs in the language analysis. The major research questions are:

1. Can LLMs identify statements that reflect burnout in persons employed in occupations with occupational stress?
2. Can LLMs identify statements that reflect depression in persons employed in occupations with occupational stress?
3. What is the proportion of burnout compared to the proportion of depression statements within and across all participants?

Results

For each transcript, both LLMs were prompted to rate participants' spoken segments as indicating burnout, depression, both, or neither. For participants, *neither* was the most frequently chosen rating; statements rated as *neither* were factual comments, brief acknowledgments, professional discussions, and yes/no responses to interviewer questions and prompts. *Both* ratings were very infrequent, tallying only 8 ratings across both LLMs and all participants. The LLMs agreed on the *neither* and burnout categories most consistently. The primary source of disagreement was ChatGPT coding 53 statements as reflecting burnout that Claude coded as *neither*, and 22 statements coded in the reverse direction. Based on a total of 1,010 pair ratings for the 11 participants, kappa equaled .658, indicating substantial agreement between the LLM ratings ^[30].

Table 1 displays data regarding the LLM ratings of burnout and depression ratings per participant. The LLMs rated speech turns from 11 participants (57,659 words), and the total number of statements averaged 92.45 ($SD = 13.93$, range = 60 to 113) per participant. The number of ChatGPT-rated burnout segments averaged 18% of the total number of segments, while ClaudeAI-rated burnout segments averaged 12% of the total. The number of segments was not significantly correlated with the number of burnout or depression ratings. The number of ChatGPT and ClaudeAI burnout ratings was significantly correlated ($r = .72$, $p = .01$); the number of burnout and depression ratings were not significantly correlated for either AI program ($r = -.18$, $p = .60$, for ChatGPT and $r = .02$, $p = .95$, for ClaudeAI).

The final row in Table 1 displays means, standard deviations, and the ratio of standard deviation divided by the mean. A ratio between .05 and .15 indicates that the scale has enough variation to be useful for statistical analyses ^[31]. The total number of participant segments falls in this range, while other indices exhibit higher variability. At the participant level, ChatGPT-rated burnout segments ranged from 40% of the total segments (Participant 6) to 4% (Participant 9). ClaudeAI-rated burnout segments ranged from 22% (Participants 10 and 11) to 4% (Participant 9). Only Participants 3 and 4 produced depression segments. ClaudeAI accounted for most of the depression ratings, while ChatGPT rated more segments as indicating burnout than ClaudeAI.

Discussion

With a substantial degree of agreement indicated by a κ of 0.658, LLM analyses offer a complementary perspective on the longstanding controversy involving the sizeable correlations found between scores on

quantitative measures of burnout and depression. The LLMs were able to identify (a) both burnout and depression statements and (b) whether a particular statement reflected burnout or depression.

The most noteworthy result was the observed individual differences in research participants regarding their expressions of burnout and depression. The assumptions of quantitative research are typically nomothetic in nature: researchers who analyze total scores from burnout and depression self-report measures at least implicitly assume that employees experience, express, and respond to occupational stress in similar ways. Two employees with a high score on a burnout scale, for example, are expected to be experiencing the same level of occupational stress. In this study, however, the relationship between burnout and depression appeared to be idiographic, meaning that the burnout/depression relationship differed by individual. Different healthcare workers expressed (a) both burnout and depression and (b) burnout but no depression. This variation in the burnout/depression relationship fits well with other research that documents changes in this correlation by age and work experience ^[1].

One possible explanation for the relative absence of depression ratings in this sample is that these workers had not yet experienced sufficient occupational stress exposure for their distress to develop into more generalized depression. Study participants reported three or fewer years of work experience. In a dose-response model ^[32], once burnout reaches a sufficient threshold of intensity and duration, it may generalize to depression. Research indicates that depression risk increases with higher intensity or repeated exposure to job strain ^[33], and psychosocial work factors such as high demands, low control, and work-family conflict also predict depressive symptoms in proportion to their frequency of exposure ^[34]. Quantitative workload indicators, such as long working hours, exhibit similar graded risk patterns ^[32].

Regarding study limitations, non-verbal communication often provides different and sometimes more valid expressions of psychosocial characteristics ^[19]. Consequently, future research should compare analysis of text and nonverbal information for evaluating occupational stress and burnout. Additionally, archival transcripts in research present both strengths and limitations: reactivity is minimized because participants are unaware of research questions, but additional questions that arise during the research process cannot be addressed. Even though analysis of interview transcripts removes many issues present in self-report measures, retrospective bias, differences in willingness to self-disclose, and interviewer effects remain.

Finally, the ability of LLMs to process sizeable amounts of text means that future research should replicate and extend these initial findings with larger language, diverse, and occupational samples.

<i>Participant# (Gender, age)</i>	<i>Total # of Participant Segments</i>	<i>#ChatGPT Burnout Ratings (% Total)</i>	<i># Claude Burnout Ratings (% Total)</i>	<i># ChatGPT Depress Ratings (% Total)</i>	<i># Claude Depress Ratings (% Total)</i>
1 (M, 32)	95	12 (13%)	6 (6%)	0	0
2 (F, 26)	90	17 (19%)	8 (9%)	0	0
3 (F, 27)	88	22 (25%)	15 (17%)	0	6 (7%)
4 (F, 37)	97	12 (12%)	6 (6%)	6 (6%)	4 (4%)
5 (F, 31)	106	6 (6%)	6 (6%)	0	0
6 (F, 24)	81	32 (40%)	15 (19%)	0	0
7 (F, 44)	113	21 (19%)	14 (12%)	0	0
8 (M, 30)	60	13 (22%)	4 (7%)	0	0
9 (F, 24)	89	4 (4%)	4 (4%)	0	0
10 (F, 36)	99	21 (21%)	22 (22%)	0	0
11 (F, 36)	99	20 (20%)	22 (22%)	0	0
Totals	1,017	180 (18%)	122 (12%)	6 (1%)	10 (1%)
<i>M</i>	92.45	16.36	11.09	0.55	0.91
<i>SD</i>	13.93	8.02	6.82	1.81	2.07
<i>SD / M</i>	.15	.49	.61	3.29	2.27

Table 1. AI Classification of Healthcare Workers' Language Samples ^[23]

Note. Participant# refers to a particular healthcare worker, while Total # of Participant Segments indicates the number of separate participant statements during the interview. The LLMs were prompted to rate participants'

speech turns as indicating Burnout, Depression, Both, or Neither. The respective columns report the number of rated Burnout and Depression entries, with percentages of the total number of segments in parentheses.

Appendix A

AI Prompt for Detecting Burnout and Depression in Text

I will upload a transcript from [transcript description]. I would like you to examine [text] and identify it as reflecting burnout, depression, both burnout and depression, or neither. Place each quote designation, your classification, and the rationale in an APA formatted table.

The language of depressed individuals tends to be focused inward, hopeless, self-critical, global, and heavy, while the language of burned-out individuals tends to be focused outward, exhausted, frustrated, cynical, and specific to work/stress context.

Here are additional examples of burnout and depression expressed verbally:

- **Depressed individual:**
 - More likely to use *negative emotional words* (“hopeless,” “worthless,” “guilty”).
 - Stronger focus on sadness, emptiness, or self-criticism.
 - Emotional language often feels heavy, global, and self-directed.
- **Burned-out individual:**
 - May use language of *exhaustion* and *frustration* (“drained,” “can’t keep up,” “overwhelmed”).
 - Emotional tone is often more about *irritability, cynicism, and stress* than despair.
 - Complaints often point outward (toward work, colleagues, or the system).
- **Depressed individual:**
 - More likely to use *negative emotional words* (“hopeless,” “worthless,” “guilty”).
 - Stronger focus on sadness, emptiness, or self-criticism.
 - Emotional language often feels heavy, global, and self-directed.
- **Depressed individual:**
 - Language is self-focused: higher use of “I” pronouns.
 - Rumination about the past and general negative worldview.
 - Generalized hopelessness (“things will never get better”).
- **Burned-out individual:**
 - Language often work- or task-focused.

- Complaints about external demands, workload, bureaucracy, or lack of support.
- Less about self-worth, more about capacity and resources.
- **Depressed individual:**
 - Absolute or all-or-nothing phrasing (“always,” “never”).
 - Catastrophic framing (“nothing matters,” “I can’t go on”).
 - Slowed, sometimes flat or sparse speech, reflecting psychomotor slowing.
- **Burned-out individual:**
 - Pragmatic, detail-oriented complaints (“too many meetings,” “I can’t finish deadlines”).
 - Language often indicates mental overload but not necessarily hopelessness.
 - May still show energy in venting, but in a cynical or detached way.
- **Depressed individual:**
 - Withdrawal signals: “I don’t want to see anyone,” “I’m a burden.”
 - Loss of interest and pleasure in social engagement.
- **Burned-out individual:**
 - Social talk may turn sarcastic, bitter, or resentful, especially about colleagues or supervisors.
 - Can still maintain some humor or venting with peers.

References

1. ^a ^b Author names omitted for ongoing review. (2026). Title withheld.
2. ^Δ Glass DC, McKnight JD (1996). "Perceived Control, Depressive Symptomatology, and Professional Burnout: A Review of the Evidence." *Psychol Health*. 11:23–48. doi:[10.1080/08870449608401975](https://doi.org/10.1080/08870449608401975).
3. ^a ^b ^c ^d ^e Bianchi R, Schonfeld IS, Laurent E (2015). "Burnout–Depression Overlap: A Review." *Clin Psychol R ev*. 36:28–41. doi:[10.1016/j.cpr.2015.01.004](https://doi.org/10.1016/j.cpr.2015.01.004).
4. ^a ^b ^c ^d Koutsimani P, Montgomery A, Georganta K (2019). "The Relationship Between Burnout, Depression, and Anxiety: A Systematic Review and Meta-Analysis." *Front Psychol*. 10:284. doi:[10.3389/fpsyg.2019.00284](https://doi.org/10.3389/fpsyg.2019.00284).
5. ^Δ Roach KN, O'Brien KE, Beehr TA, Glazer S (2025). "A Construct of Its Own: Examining the Overlap of Burnout with Depression, Anxiety, and Fatigue." *Occup Health Sci*. 9:1–40. doi:[10.1007/s41542-025-00224-3](https://doi.org/10.1007/s41542-025-00224-3).
6. ^a ^b ^c Schonfeld IS, Bianchi R (2022). "Distress in the Workplace: Characterizing the Relationship of Burnout Measures to the Occupational Depression Inventory." *Int J Stress Manag*. 29(3):253–259. doi:[10.1037/str0000261](https://doi.org/10.1037/str0000261).

7. [△]Bianchi R, Schonfeld IS (2020). "The Occupational Depression Inventory: A New Tool for Clinicians and Epidemiologists." *J Psychosom Res.* **138**:110249. doi:[10.1016/j.jpsychores.2020.110249](https://doi.org/10.1016/j.jpsychores.2020.110249).
8. [△][△]Tavella G, Hadzi-Pavlovic D, Parker G (2021). "Burnout: Redefining Its Key Symptoms." *Psychiatry Res.* **302**:114023. doi:[10.1016/j.psychres.2021.114023](https://doi.org/10.1016/j.psychres.2021.114023).
9. [△]Tavella G, Parker G (2020). "A Qualitative Reexamination of the Key Features of Burnout." *J Nerv Ment Dis.* **208**(6):452–458. doi:[10.1097/NMD.0000000000001155](https://doi.org/10.1097/NMD.0000000000001155).
10. [△]Kalani S, Karimi Z, Ahmadi M (2025). "Is the "Burnout Epidemic" an Academic Fiction? A Systematic Review of Tools for Estimating the Prevalence and Assessment of Burnout and Introduction of New Instruments." *J Psychol.* **29**(2):187–200. doi:[20.1001.1.18808436.1404.29.2.177](https://doi.org/20.1001.1.18808436.1404.29.2.177).
11. [△]Schwarz N (2010). "Measurement as Cooperative Communication: What Research Participants Learn from Questionnaires." In: Walford G, Tucker E, Viswanathan M, editors. *The SAGE Handbook of Measurement*. Sage Publications. p. 43–60.
12. [△][△]Dohrenwend BP (2006). "Inventorying Stressful Life Events as Risk Factors for Psychopathology: Toward Resolution of the Problem of Intracategory Variability." *Psychol Bull.* **132**(3):477–495. doi:[10.1037/0033-2909.132.3.477](https://doi.org/10.1037/0033-2909.132.3.477).
13. [△]Bayerl SP, Roccabruna G, Chowdhury SA, Ciulli T, Danieli M, Riedhammer K, et al. (2022). "What Can Speech and Language Tell Us About the Working Alliance in Psychotherapy?" In: *Proceedings of Interspeech 2022*. p. 347–351. doi:[10.21437/Interspeech.2022-347](https://doi.org/10.21437/Interspeech.2022-347).
14. [△][△]Shapira N, Lazarus G, Goldberg Y, Gilboa-Schechtman E, Tuval-Mashiach R, Juravski D, et al. (2021). "Using Computerized Text Analysis to Examine Associations Between Linguistic Features and Clients' Distress During Psychotherapy." *J Couns Psychol.* **68**(1):77–87. doi:[10.1037/cou0000440](https://doi.org/10.1037/cou0000440).
15. [△]Zilcha-Mano S, Krasovsky T (2024). "Using Virtual Reality to Understand Mechanisms of Therapeutic Change." *Nat Rev Psychol.* Advance online publication. doi:[10.1038/s44159-024-00303-2](https://doi.org/10.1038/s44159-024-00303-2).
16. [△][△]Author names omitted for ongoing review. (2026). Title withheld.
17. [△]Shostrum EL, producer (1966). *Three Approaches to Psychotherapy* [Film]. Santa Ana, CA: Psychological Films.
18. [△]Rathje S, et al. [Details masked for peer review — downloaded 7 Feb, 2024]
19. [△][△]Author names omitted for ongoing review. (2026). Title withheld.
20. [△]Grisso MK, Springstein T (2025). "Generalizing Across People or Capturing Unique Personal Experiences? Nomothetic to Idiographic Ways of Assessing Self-Reported Emotions in Daily Life." *Transl Issues Psychol Sci.* Advance online publication. doi:[10.1037/tps0000484](https://doi.org/10.1037/tps0000484).

21. ^a ^bFullen RN (2025). "Perceptions of Clinically Depressed Individuals: Social Media Use, Depressive Symptoms, and Interpersonal Relationships." *Walden Dissertations and Doctoral Studies*. 17363. <https://scholarworks.waldenu.edu/dissertations/17363>.
22. ^a ^bLi Y, Xiang D, Jiang Y, Gu S (2025). "Understanding Burnout Among Operating Room Nurses: A Qualitative Study." *Front Public Health*. 13:1604631. doi:[10.3389/fpubh.2025.1604631](https://doi.org/10.3389/fpubh.2025.1604631).
23. ^a ^b ^cSmith MA, Hoult LM, Rippon D, O'Brien N, Branley-Bell D, Byrne-Davis L, et al. (2025). "Facilitators and Barriers to Engaging in Expressive Writing Among Health and Social Care Professionals." *PLoS ONE*. 20:e0328801. doi:[10.1371/journal.pone.0328801](https://doi.org/10.1371/journal.pone.0328801).
24. ^aAuthor names omitted for ongoing review. (2026). Title withheld.
25. ^aEdwards T, Holtzman NS, Eshleman A (2019). "A Meta-Analysis of Depression and Self-Focused Language." *J Res Pers*. 78:151–166. doi:[10.1016/j.jrp.2018.11.005](https://doi.org/10.1016/j.jrp.2018.11.005).
26. ^aMaslach C, Leiter MP (2016). "Understanding the Burnout Experience: Recent Research and Its Implications for Psychiatry." *World Psychiatry*. 15(2):103–111. doi:[10.1002/wps.20311](https://doi.org/10.1002/wps.20311).
27. ^aRude SS, Gortner EM, Pennebaker JW (2004). "Language Use of Depressed and Depression-Vulnerable College Students." *Cogn Emot*. 18(8):1121–1133. doi:[10.1080/02699930441000030](https://doi.org/10.1080/02699930441000030).
28. ^aToker S, Biron M (2012). "Job Burnout and Depression: Unraveling Their Temporal Relationship and Considering the Role of Physical Activity." *J Appl Psychol*. 97(3):699–710. doi:[10.1037/a0026914](https://doi.org/10.1037/a0026914).
29. ^aOpenAI (2024). "ChatGPT: Language Model for Natural Language Understanding and Generation." Open AI. <https://openai.com>.
30. ^aLandis JR, Koch GG (1977). "The Measurement of Observer Agreement for Categorical Data." *Biometrics*. 33(1):159–174. doi:[10.2307/2529310](https://doi.org/10.2307/2529310).
31. ^aDavis CS (1987). *Statistical Methods for the Analysis of Repeated Measurements*. New York: Springer.
32. ^a ^bRugulies R, Aust B, Burr H, Bültmann U, Nielsen MBD (2013). "Effort–Reward Imbalance at Work and Risk of Severe Depressive Symptoms in the Danish Workforce." *Eur J Public Health*. 23:617–622. doi:[10.1093/eurpub/cks124](https://doi.org/10.1093/eurpub/cks124).
33. ^aStansfeld SA, Shipley MJ, Head J, Fuhrer R (2012). "Repeated Job Strain and the Risk of Depression: Longitudinal Analyses From the Whitehall II Study." *Am J Public Health*. 102(12):2360–2366. doi:[10.2105/AJPH.2011.300589](https://doi.org/10.2105/AJPH.2011.300589).
34. ^aNiedhammer I, Bertrais S, Witt K (2015). "Psychosocial Work Exposures and Health Outcomes: A Study Among Employees in France." *BMC Public Health*. 15:93. doi:[10.1186/s12889-015-1440-9](https://doi.org/10.1186/s12889-015-1440-9).

Declarations

Funding: No specific funding was received for this work.

Potential competing interests: No potential competing interests to declare.