Psychology & Health

ISSN: 0887-0446 (Print) 1476-8321 (Online) Journal homepage: www.tandfonline.com/journals/gpsh20

Associations between burnout, employee silence and voice: a systematic review and meta-analysis

Olga Lainidi, Judith Johnson, Bethany Griffin, Panagiota Koutsimani, Christos Mouratidis, Chris Keyworth & Daryl B. O'Connor

To cite this article: Olga Lainidi, Judith Johnson, Bethany Griffin, Panagiota Koutsimani, Christos Mouratidis, Chris Keyworth & Daryl B. O'Connor (28 May 2025): Associations between burnout, employee silence and voice: a systematic review and meta-analysis, Psychology & Health, DOI: 10.1080/08870446.2025.2509074

To link to this article: https://doi.org/10.1080/08870446.2025.2509074

PSYCHOLOGY & HEALTH
https://doi.org/10.1080/08870446.2025.2509074

Associations between burnout, employee silence and voice: a systematic review and meta-analysis

Olga Lainidi[a], Judith Johnson[b], Bethany Griffin[a], Panagiota Koutsimani[c], Christos Mouratidis[d], Chris Keyworth[a] and Daryl B. O'Connor[a]

[a]School of Psychology, Faculty of Medicine and Health, University of Leeds, Leeds, UK; [b]Division of Nursing, Midwifery and Social Work, University of Manchester, Manchester, UK; [c]Department of Psychology, Aristotle University of Thessaloniki, Thessaloniki, Greece; [d]School of Psychology, University of Derby, Derby, UK

ABSTRACT

Objective: The intention to speak-up or withhold one's voice is linked to employee well-being outcomes and is considered a proxy for the quality of organisational culture in the workplace. This systematic review and meta-analysis synthesised evidence on the relationship between burnout and employee silence/voice outcomes.

Methods: An electronic database search up to May 2024 was conducted on eight databases combined with manual scoping of references and 84 studies met the inclusion criteria ($N = 34{,}975$).

Results: The relationship between all employee voice/silence outcomes and burnout was statistically significant with greater silence and lower voice being associated with higher burnout ($\rho = .36$, 95% *CI* [.32, .40]). Examined separately, effects were moderate and positive for silence and burnout ($\rho = .43$, 95% *CI* [.37, .48]) and small and negative for voice and burnout ($\rho = -0.28$, 95% *CI* [–0.35, –0.21]). Subgroup analyses revealed larger effects in non-Western regions and studies using the Maslach-Burnout-Inventory.

Conclusions: The evidence consistently showed a larger overlap between burnout and silence, compared to voice, suggesting that reducing silence is more beneficial for addressing burnout than increasing voice. The evidence is limited primarily to emotional exhaustion, and more research is needed to distinguish the emotional/cognitive components of silence/voice from behavioural outcomes.

ARTICLE HISTORY
Received 15 November 2024
Accepted 8 May 2025

KEYWORDS
Burnout; employee silence; employee voice; meta-analysis

Introduction

Employee silence and employee voice

The recent emphasis on fostering work environments where employees feel empowered to express their thoughts, opinions and concerns consistently reveals that workers often hesitate to express their opinions about different issues regarding their job and/or their organisation (Donaghey et al., 2019; Morrison, 2023). This is often referred

CONTACT Olga Lainidi o.lainidi@leeds.ac.uk School of Psychology, Faculty of Medicine and Health, University of Leeds, Leeds LS2 9JT, UK.

Supplemental data for this article can be accessed online at https://doi.org/10.1080/08870446.2025.2509074.

to as employee silence and has been defined as 'the intentional withholding of any form of genuine expression about the individual's behavioural, cognitive, and/or affective evaluations of their circumstances to persons who are perceived to be capable of effecting change or redress' (Pinder & Harlos, 2001, p. 334). Employee voice, on the other hand is defined as the informal and discretionary communication of employees' evaluations of personal, social, and/or organisational circumstances at work to persons who are capable of affecting change (Morrison, 2023) and is often regarded as a helping or an extra-role behaviour (LePine & Van Dyne, 1998).

Earlier academic work described voice as a response to dissatisfaction when employees are not happy with the current state of affairs in the organisation (e.g. Hirschman, 1970). More recently, the Organizational Behavior literature has focused on employee voice as a positive behavior, indicative of active engagement and organisational citizenship (e.g. LePine & Van Dyne, 1998, voice as an extra-role behavior). Extending this work, Van Dyne et al. (2003) proposed three forms of employee silence and employee voice behaviour based on the motive they reflect: 1) quiescent/defensive silence or voice reflecting a self-protective behaviour (Knoll & van Dick, 2013; Van Dyne et al., 2003); 2) acquiescent silence or voice reflecting disengaged behaviour due to feeling that there is no point to put effort in changing things (Morrison & Milliken, 2000) and 3) prosocial silence or voice reflecting an other-oriented motive to protect the organisation, primarily linked to Organizational Citizenship Behavior (OCB) theory (Mowbray et al., 2021). To these three motives, Knoll and van Dick (2013) added a fourth motive for silence, that of opportunistic silence, which is associated with selfish motives aiming to ensure privileges for oneself (Knoll et al., 2019); opportunistic silence is generally seen as a counterproductive work behavior assumed to be motivated by deviance (Connelly et al., 2012). Opportunistic silence is rooted in knowledge hiding, which refers to '…an intentional attempt by an individual to withhold or conceal knowledge that has been requested by another person…' (Connelly et al., 2012. p.12), thus expanding the definition of employee silence to include more strategic forms.

In terms of understanding employee voice, Liang et al. (2012) proposed two forms, promotive and prohibitive voice with the former based on employees' intention to improve organisational functioning (similar to LePine & Van Dyne, 1998) and the latter to communicate concerns about problems that can lead to negative consequences (similar to quiescent/defensive voice). Other types of voice in the literature include voice based on self-oriented motives, such as aggressive voice (Hagedoorn et al., 1999) where employees continually confront others until they achieve desirable results and self-interested voice, focused on the personal benefits of the individual employee (Duan et al., 2021) – similar to opportunistic silence and knowledge hiding behavior.

Silence, voice and burnout

While speaking up and withholding concerns have an inverse relationship with each other conceptually, the weak correlation between the two ($\rho = -.15$; Sherf et al., 2021) suggests that silence and voice can operate independently of each other, meaning that employees might be speaking-up about certain things but remaining silent about others. The limited overlap indicates that speaking up more does not necessarily mean employees are withholding equally less and vice versa. The available evidence

points towards employee voice having much stronger relationships with measures of work autonomy, while employee silence has a stronger relationship with negative affect and perceived psychological safety (Hao et al., 2022). It is not yet clear whether this could be partially explained by methodological artifacts of the measures used for voice and silence, whether this could be the result of participants pooling from different experiences regarding voice and silence, or a mix of the two. One way to further our understanding of the two is *via* delineating their relationships with key indicators of work-related well-being such as burnout.

Burnout is a psychological syndrome involving a prolonged response to chronic emotional and interpersonal stressors on the job (Maslach & Leiter, 2006) and is arguably, one of the most researched constructs in occupational health in spite of the lack of an agreed upon definition (Schaufeli et al., 2023). The most widely used approach to burnout views it as a multidimensional construct consisting of three key dimensions: exhaustion, feelings of cynicism/depersonalisation, and a sense of reduced professional efficacy/lack of accomplishment as measured with the Maslach Burnout Inventory (MBI) (Leiter & Maslach, 2016). Unidimensional approaches to burnout usually reduce it to mainly exhaustion due to work-related problems, whereas Schaufeli and Taris (2005) position a combined lack of energy with mental distancing from work at its core. This has implications for the study of burnout and the statistical relationships identified with other constructs. More recently, burnout has been enriched by the addition of cognitive impairment as a component (Schaufeli et al., 2020) based on accumulated research evidence of both self-reported and objectively assessed decline in cognitive functions (Koutsimani & Montgomery, 2022; Renaud & Lacroix, 2023) and supported by the Conservation of Resources (COR) theory (Hobfoll & Freedy, 2017). This updated conceptualisation has resulted in the recently developed Burnout Assessment Tool (BAT), which operationalises burnout as a syndrome comprising exhaustion, mental distance, and symptoms of both cognitive and emotional impairment (Schaufeli et al., 2020).

There are two existing meta-analyses that have reported on the relationships between employee silence/voice and burnout, one by Hao et al. (2022) and one by Sherf et al. (2021). Both reported a stronger relationship between greater silence and greater burnout ($\rho = .40$; Hao et al., 2022, $\rho = .32$; Sherf et al., 2021) compared to the relationship between greater employee voice and less burnout ($\rho = -0.11$; Sherf et al., 2021). Existing attempts to explain these relationships provide an overarching and descriptive account of the differences in antecedents/consequences of silence and voice (e.g. informed by the Behavioral Activation and Inhibition systems, BIS and BAS). The available evidence suggests that voice shows a marginal negative relationship with negative affect ($\rho = -.09$) and a comparatively stronger one with positive affect ($\rho = .20$; Chamberlin et al. 2017); silence, on the other hand, shows stronger relationships with both negative affect ($\rho = .26$) and positive affect ($\rho = -0.17$; Hao et al., 2022), indicating that silence might have a wider emotional component compared to voice. Similarly, emotional exhaustion and cynicism/depersonalisation have equally strong but opposing relationships with negative and positive affectivity (Alarcon et al., 2009), suggesting parallels between silence and the two core burnout components. Thus, a more consistent relationship between silence and emotional exhaustion as well as silence and cynicism/depersonalisation is expected, supported by work–stress theories,

such as the COR theory (Hobfoll & Shirom, 2000) and the Job Demands-Resources model (JD-R; Bakker & Demerouti, 2007). Given that even positive types of silence are linked to greater burnout (e.g. prosocial silence), it can be assumed that the subjective experience of withholding any form of genuine expression is either more common among emotionally exhausted employees or requires high levels of emotional and cognitive self-regulation, (e.g. *via* engagement in an ongoing suppression involving rumination both in and outside of work thus also consuming recovery resources; Sonnentag & Bayer, 2005) – or a combination of both. Delineating the relationships between voice and burnout is more complicated due to its heterogeneity and depends on how voice is operationalised, as certain types of voice are expected to be positively linked to burnout (e.g. voice to express employee dissatisfaction with the current state of affairs in the organisation; Hirschman, 1970) and others negatively (e.g. voice as a positive, extra role behavior; LePine & Van Dyne, 1998).

The current meta-analysis addresses important gaps in the literature. First, research on this topic area has increased exponentially since 2020 when the most recent meta-analyses on these topics completed their searches, and it is possible that the significantly larger amount of research data could lead to more nuanced results. Second, it is currently unclear whether different operationalisations of employee silence lead to different associations with burnout. The previous meta-analyses focused mainly on employee silence motives or grouped these factors together preventing a more detailed understanding to be gained. The current systematic review and meta-analysis examines the different operationalisations of employee silence, beyond silence motives, including studies that measured knowledge-hiding behaviors, studies that asked the participants about silence content (e.g. withholding information about a solution to a problem regardless of the underlying reason), as well as studies where participants were asked about the norms of voice/silence in their organisation (voice climate).

Third, it is currently unknown whether different dimensions of burnout have different relationships with voice and silence. Burnout is widely recognised as a tripartite construct, combining exhaustion, cynicism and professional efficacy, but the relationships of silence and voice with each dimension separately have not been analysed in the previous meta-analyses.

Fourth, it is unknown whether the associations between burnout, voice and silence are moderated by study level factors such as sample-level (e.g. % female, mean age of participants), study-level (e.g. design, response rate, geographical region) and measure level (e.g. reliability coefficient and number of items for burnout measure) characteristics. The current meta-analysis tested a wider array of sample-level, study-level and measure-level moderators to explore potential influencing factors.

Specific objectives

This systematic review and meta-analysis aimed to summarise and clarify the existing evidence concerning the relationship between burnout and employee silence and voice outcomes (Supp. Material 1. PICO). There were three objectives:

1. To examine whether burnout is associated with employee silence outcomes (e.g. frequency of reported silence behaviours, employee silence beliefs);

2. To examine whether burnout is associated with employee voice outcomes (e.g. frequency of reported voice behaviours, employee voice beliefs)
3. To examine whether the associations between burnout and employee silence outcomes and employee voice outcomes are moderated by study-level, sample-level and measure-level factors (e.g. gender of participants, geographical area, % female participants).[1]

Method

Search strategy

A review protocol was registered on PROSPERO (CRD42023384630). Initially, an electronic database search with no time restriction was conducted covering all years until May 2023 which then was updated with a second search to May 2024. The following electronic databases and search engines were searched: PubMed, PsycINFO, Scopus, Embase, Cochrane Library, Web of Science, CINAHL and Google Scholar. We also consulted the reference lists of major reviews (Morrison, 2023) and existing meta-analyses (Hao et al., 2022; Sherf et al., 2021). The keywords identified were then put together into a search string adjusted for each database (Supplemental Material 1). The search included both title and abstract.

The following inclusion criteria were applied: 1) studies using quantitative methodology (cross-sectional and longitudinal studies), 2) with samples comprising employed individuals only, 3) including a measure of burnout (e.g. Maslach Burnout Inventory, Burnout Assessment Tool, Oldenburg Burnout Inventory) and a measure of employee silence (e.g. Employee Silence Scale by Pinder & Harlos, 2001) or a measure of employee voice (e.g. Promotive Voice Scale by Liang et al., 2012), 4) published peer-reviewed research articles, 5) where full-text articles could be retrieved and 6) were written in the English language. Relevant papers that were identified as grey literature were excluded. Where we could not find the items within a measure or the measure was not specified, we excluded the paper from the review. Prior to excluding studies where the measure of silence/voice was not specified or could not be retrieved, we emailed the authors twice asking for that information to be provided. Finally, where coefficients for the meta-analysis were not provided in the manuscript, authors were contacted twice in an attempt to obtain this information.

The initial search of databases yielded 1733 results. The data retrieved from each database was logged in Rayyan where duplicate control was conducted, resulting in 1509 citations to screen. Two reviewers (OL and PK) independently screened approximately 10% of all identified abstracts with a satisfactory agreement (Cohen's kappa = .92, 96.4% agreement), and the remainder was screened by one reviewer (OL). Two reviewers (OL and PK) independently screened all articles retained for full-text screening; the patterns of the disagreement indicated conflicting decisions in including studies reporting on organisational/job cynicism and knowledge hiding. The disagreement was resolved by examining the measures used for each paper to align with the inclusion/exclusion criteria; papers were included when cynicism was measured as a burnout component (e.g. MBI cynicism) and when knowledge hiding measures included at least one subscale fitting the opportunistic silence description. The selection

procedure is presented in a Preferred Reporting Items for Systematic Reviews and Meta-analyses (PRISMA) flowchart (Supplemental Material 1, Figure 1). A total of 508 papers were retrieved for full-text screening. Wrong outcome was the most common reason for exclusion at the full-text screening stage ($n = 363$ papers) (Supplemental Material 1, Figure 1 Reasons for Exclusion). A total of 101 full-text papers were included at the data extraction stage; however, 20 papers were subsequently excluded due to unavailable correlation coefficients (or other coefficients that could be converted) after emailing the respective authors twice.

Quality appraisal

Quality assessment was conducted using the Quality Assessment Tool for Observational Cohort and Cross-Sectional Studies (Feng et al., 2014). Two reviewers (OL and CM) independently assessed the quality of the included studies. The tool contains 14 criteria, and the evaluator is asked to answer whether the study in question meets the criterion, with the possible answers being 'Yes, No, Cannot Determine, Not applicable, and Not Reported'. A score of >10 corresponds to good quality, 7–10 to fair quality and <7 to poor quality. Initial agreement was low (Cohen's kappa = .32, 63.2%) due to inconsistent interpretation of criteria 2, 3, 7 and 14 between the two reviewers. The clarification and agreement on the application of these four criteria resulted in a substantial improvement in the reviewers' agreement (Cohen's kappa = .79, 90.03%).

Data coding procedure

All collected data was extracted in a Microsoft Excel document where the following descriptive information of each study was included: study setting details (lead author and date, study location/country, sampling, industry, study design), participants' characteristics (sample size at baseline and follow-up were applicable, age mean/SD or median/range, female %, occupational group, industry, working experience mean/SD or median/range) and burnout and silence or voice measure(s) (measure of burnout, measure of employee silence and/or voice outcome, reliability coefficients, number of items for burnout measure), and correlation coefficients for the employee silence/employee voice burnout relationship. In terms of the measures of voice and silence, we followed a coding approach similar to the meta-analysis of Sherf et al. (2021) by reviewing labels and items measuring each construct. All employee silence and knowledge hiding outcomes were coded as 'employee silence'; where possible, subcategories were coded based on the measures used including acquiescent, quiescent, prosocial, self-interested silence (includes opportunistic silence and knowledge hiding) and silence content (the topic about which one is being silent). Employee voice outcomes reporting on the person's behavior/intention or motivation were coded as 'employee voice'; due to the limited number of studies identifying different forms of voice (e.g., aggressive or self-oriented voice), coding subcategories was not meaningful. Measures asking participants to report their views on the extent to which their work environment encourages or discourages voice or silence were coded as 'voice climate' (Supplemental Material 2, Table with included studies). Two authors (OL and BG) independently coded 10% of the included studies with acceptable agreement rate (91.2%) and the remainder was coded by one reviewer (OL).

To conduct a meta-analysis across both outcomes, we first standardised the direction of the effects by reverse scoring the voice coefficients. More specifically, for the overall meta-analysis, all studies were coded such that stronger coefficients indicating greater burnout were associated with greater silence/less voice. To further decompose the main meta-analysis, for the next step, all employee silence outcomes were manually coded as 'negative', (i.e. greater burnout was associated with greater silence) and all employee voice outcomes as well as voice climate were coded as 'positive' (i.e. greater burnout was associated with less voice and vice versa) (Supplemental Material 3)[2]. This allowed us to test for the absolute strength of the relationships under the assumption that greater silence is associated with a more negative work environment, whereas greater voice is associated with a more positive one. This also applied to studies where silence/voice outcomes were rated by a supervisor/colleague to simulate the direction of the effects in the self-reported studies.

Data analysis

A meta-analysis was conducted using the Comprehensive Meta-Analysis software (version 3) (CMA; Borenstein et al., 2009; Lipsey & Wilson, 2001). Specifically, the CMA software uses a Fisher's *Z*-transformation to weight the correlation coefficients by the reciprocal quantity of the error variance and decrease the bias further. We adopted a random-effects model and a weighted correlation coefficient (ρ) with 95% confidence intervals (95% *CI*) was used as an overall synthesised measure of effect size. For studies reporting more than one relevant correlation coefficient for the same sample, a weighted correlation coefficient was calculated, whereas in studies reporting effect sizes from independent samples, each sample was included as a separate study. Forest plots were constructed to visually represent the mean effect sizes. The *Q* within-statistic was used to assess the heterogeneity of studies which tests the null hypothesis that all studies in an analysis share a common effect size (Borenstein et al., 2009). Following the recent proposition by Borenstein (2022) that the I^2 statistic is not a quantifier of heterogeneity, this statistic was not reported. 'Leave-one-out' sensitivity analysis was used to assess the robustness of the synthesised results.

Subgroup analyses using mixed-effects analysis were conducted to explore potential moderators that could explain heterogeneity. As a rule of thumb, subgroup analyses were probed where either a minimum number of 10 studies-per-group were available (Higgins et al., 2024) or a minimum of 20 studies in total (Belias et al., 2019). Following Higgins et al. (2024), the use on the *Q* test in subgroup analysis was treated as merely an indicator of heterogeneity and was not considered evidence of absence/presence of statistically significant differences between the examined groups; thus, we were interested in observable differences in the effect sizes rather than the significance of the *Q* statistic (Higgins et al., 2024). Subgroup analysis was conducted for the following categorical moderators: country group (Western/Westernised vs. non-Western/Westernised), burnout measure (MBI vs. non-MBI), % female participants (> 55% vs. < 55%), whether silence/voice were measured at the same time point as burnout, quality assessment score (fair vs. poor) and sample size (N >/= 200 vs. $N < 200$). A series of meta-regression analyses were conducted where more than 10 studies were available testing the following

continuous moderators: response rate, publication year, female percentage (continuous), sample size, mean age, mean years in the organisation, number of items in burnout measure, reliability estimates for burnout and quality assessment score (continuous).

Three indicators of publication bias were examined: standard error funnel plots for observed only and imputed studies, Duval and Tweedie's trim and fill procedure and Egger's regression intercept (Duval & Tweedie, 2000; Egger et al., 1997; Lipsey & Wilson, 2001; Sedgwick, 2013). As Rosenthal's method has been criticised that it fails to account for the bias in the 'file drawer' of unpublished studies, and thus can give misleading results (Scargle, 1999), we calculated Orwin's fail-safe N. The original coefficients extracted from the studies were used to test for publication bias.

Results

Study characteristics

In total, 84 studies (81 research articles) were included in this review (Supplemental Material 2, Table with included studies). The total sample of participants was 34,975 ($N = 34{,}975$, 48.22% identified as female, 11 studies did not provide data on gender). Twenty-two studies included a sample of various employees (26.2%), followed by health and social care (14 studies, 16.7%) and education (9 studies, 10.7%). Fifty-nine (70.2%) of the included studies were conducted in non-Western/Westernised countries, with 27 (32.9%) studies having been conducted in China. Sample sizes ranged from $N = 72$ to $N = 3{,}266$. All the studies were correlational (i.e. no experimental designs were identified), and the majority were also cross-sectional (58/84, 68.83%), with one longitudinal study and 25 studies using a time-lagged design whereby predictor variables were measured at time point 1 and outcome variables at a subsequent point, but with no baseline measurement of the outcome variables. In terms of the measures, the MBI was the most commonly used measure for burnout (67/84 studies, 79.8%), whereas 30 distinct measures were used to capture employee silence/employee voice outcomes (e.g. Van Dyne et al., 2003, employee silence motives questionnaire; LePine & VanDyne, 1998, employee voice questionnaire).

Quality appraisal

Of the 84 studies that were included in the current review, none of the included studies were rated as good, 65 as fair (77.4%), and 19 (22.6%) were rated as having poor quality (Supplemental Material 3, Quality Score and Quality Group for each study).

Meta-analysis

Main meta-analysis

The overall relationship between all employee voice/silence outcomes (coded in the same direction indicating greater burnout was associated with greater silence/less voice) and burnout was positive and statistically significant ($\rho = .36$, 95% *CI* [.32; .40], $Q = 1653.52$, $p < .001$, $k = 84$, $N = 34{,}975$). The overall effect size ranged from $\rho_{min} = .35$

to ρ_{max}=.36 when individual studies were omitted, and no single study significantly altered the overall effect. Examining employee silence and employee voice separately revealed a moderate effect size between greater employee silence and greater burnout (ρ = .43, 95% *CI* [.37; .48], $Q = 740.26$, $p < .001$, $k = 44$, $n = 13{,}633$) and a small effect size between greater employee voice and less burnout (ρ = −0.28, 95% *CI* [−0.32; −0.21], $Q = 550.35$, $p < .001$, $k = 32$, $n = 12{,}361$) and greater voice climate with less burnout (ρ = −0.29, 95% *CI* [-0.37; −0.21], $Q = 111.20$, $p < .001$, $k = 9$, $n = 9{,}062$). Greater employee silence reasons were moderately associated with greater burnout (ρ = .46, 95% *CI* [.40; .52], $Q = 538.30$, $p < .001$, $k = 35$, $n = 9{,}743$) compared to the small but significant effect size for greater employee silence content with greater burnout (ρ = .27, 95% *CI* [.22; .32], $Q = 21.06$, $p < .01$, $k = 9$, $n = 3{,}890$).[3] The Forrest Plots for each of the main meta-analyses can be found in Supplemental Material 4, Figures 2–7.

Examining each burnout component, moderate effects were observed between greater employee silence and greater emotional exhaustion (ρ = .44, 95% *CI* [.33; .54], $Q = 1502.03$, $p < .001$, $k = 33$, $n = 11{,}269$) as well as greater cynicism/depersonalisation (ρ = .37, 95% *CI* [.21; .51], $Q = 295.61$, $p < .001$, $k = 12$, $n = 3{,}696$) and a small effect between greater employee silence and reduced professional efficacy (ρ = .24, 95% *CI* [.05; .41], 134.64, $p < .001$, $k = 8$, $n = 1{,}866$). Greater employee voice was significantly associated with lower emotional exhaustion (ρ = −0.25, 95% *CI* [−0.32; −0.19], $Q = 287.37$, $p < .001$, $k = 25$, $n = 9{,}817$) as well as with greater professional efficacy (ρ = .20, 95% *CI* [−0.27; −0.13], $Q = 1.96$, $p > .05$, $k = 3$, $n = 701$) with small effect sizes, and a non-significant relationship was found between voice and cynicism/depersonalisation (ρ = −0.13, 95% *CI* [−0.27; −0.13], $Q = 8.84$, $p < .01$, $k = 2$, $n = 846$). Across studies that reported on overall burnout levels, a large effect size was observed with between higher burnout with greater employee silence (ρ = .63, 95% *CI* [.55; .69], $Q = 31.51$, $p < .001$, $k = 8$, $n = 1{,}468$) and a moderate effect size with lower employee voice (ρ = −0.39, 95% *CI* [−0.56; −0.19], $Q = 232.27$, $p < .001$, $k = 7$, $n = 3{,}353$). In terms of the different forms of employee silence, generally moderate effect sizes were observed across the two burnout dimensions, emotional exhaustion and cynicism/depersonalisation, and quiescent, acquiescent and self-interested silence, whereas smaller effect sizes were observed for reduced professional efficacy (Table 1).

Subgroup analysis and meta-regressions

As the number of available samples was $k < 20$ for studies that reported on cynicism/depersonalisation, reduced professional efficacy and total burnout scores, subgroup analyses were only probed for emotional exhaustion. Focusing on the relationship between employee silence and emotional exhaustion, the subgroup analyses revealed stronger effect sizes in studies conducted in non-Western/Westernised (ρ = .46, 95% *CI* [.34; .56]) compared to Western/Westernised countries (ρ = .33, 95% *CI* [.07; .55]), as well as in studies utilising the MBI (ρ = .46, 95% *CI* [.35; .57]), compared to non-MBI studies (ρ = .29, 95% *CI* [.002; .53]) and in studies with > 55% identifying as female (ρ = .53, 95% *CI* [.32; .70]) compared to studies with <55% female (ρ = .40, 95% *CI* [.26; .52]). Studies rated as fair showed a smaller effect size (ρ = .38, 95% *CI* [.27; .49]), compared to those rated poor (ρ = .59, 95% *CI* [.41; .72]), whereas studies with $N > 200$ had a larger effect size (ρ = .47, 95% *CI* [.35; .57]), compared to $N < 200$ (ρ = .33, 95% *CI* [.10; .53]) (Table 2).

Table 1. Main meta-analytic results for the relationships between employee silence/employee voice outcomes and burnout.

					95%CI	
	k	*N*	*ρ*	SD_ρ	[LL; UL]	*Q*
All studies (negative behaviour)	84	34,975	.36	.22	[.32; .40]	1653.52***
Employee silence	44	13,633	.43	.23	[.37; .48]	740.26***
Reasons for silence	35	9,743	.46	.23	[.40; .52]	538.30***
Silence content	9	3,890	.27	.07	[.22; .32]	21.06**
Employee voice	32	12,361	−0.28	.20	[-0.35; −0.21]	550.35***
Voice climate	9	9,062	−0.29	.12	[-0.37; −0.21]	111.20***
Emotional exhaustion						
All studies (negative behaviour)	64	28,568	.37	.28	[.30; .43]	2269.11***
Employee silence	33	11,269	.44	.37	[.33; .54]	1502.03***
Quiescent silence	11	3,466	.38	.25	[.24; .50]	206.48***
Acquiescent silence	8	2,878	.39	.30	[.20; .55]	172.34***
Prosocial silence[4]	5	1,352	.39	.51	[-0.04; .70]	269.49***
Self-interested silence	12	3,880	.41	.12	[.35; .47]	56.84***
Silence content	7	3,419	.27	.05	[.22; .32]	12.32
Employee voice (positive behaviour)	25	9,817	−0.25	.17	[−0.32; −0.19]	287.37***
Voice climate	7	8,740	−0.35	.17	[-0.46; −0.23]	194.10***
Depersonalisation/cynicism						
All studies (negative behaviour)	15	4,657	.36	.30	[.22; .49]	388.63***
Employee silence	12	3,696	.37	.30	[.21; .51]	295.61***
Quiescent silence	7	2,128	.40	.34	[.16; .59]	205.61***
Acquiescent silence	6	1,620	.44	.56	[.02 .72]	385.24***
Prosocial silence	6	1,620	.25	.24	[.05; .43]	78.26***
Self-interested silence	4	1,589	.32	.14	[.18; .44]	24.46***
Silence content	1	332	.12	.00	[.01; .23]	0.00
Employee voice (positive behaviour)	2	846	−0.13	.14	[−0.32; .07]	8.84**
Voice climate	1	115	−0.67	.00	[−0.76; −0.55]	0.00
Reduced professional efficacy						
All studies (negative behaviour)	12	2,918	.21	.23	[.07; .34]	144.22***
Employee silence	8	1,866	.24	.27	[.05; .41]	134.64***
Quiescent silence	6	1,620	.29	.25	[.09; .47]	83.8***
Acquiescent silence	6	1,620	.35	.24	[.17; .52]	76.13***
Prosocial silence [5]	6	1,620	.24	.39	[−0.07; .51]	189.60***
Self-interested silence	2	835	.26	.00	[.19; .32]	0.59
Employee voice (positive behaviour)	3	701	−0.20	.00	[−0.27; −0.13]	1.96
Voice climate	1	115	−0.04	.00	[−0.22; .14]	0.00
Total burnout score[6]						
All studies (negative behaviour)	17	5,321	.48	.29	[.37; .58]	419.66***
Employee silence	8	1,468	.63	.14	[.55; .69]	31.51***
Quiescent silence	2	279	.61	.26	[.31; .80]	8.90**
Acquiescent silence	3	457	.52	.17	[.35; .66]	7.84**
Prosocial silence	2	279	.64	.29	[.33; .83]	9.51**
Self-interested silence	1	265	.66	.00	[.59; .72]	0.00
Silence content	1	139	.40	.00	[.25; .53]	0.00
Employee voice (positive behavior)[7]	7	3,353	−0.39	.29	[-0.56; −0.19]	232.27***
Voice climate	2	298	−0.31	.00	[-0.40; −0.21]	0.15

Note. Random-effects model, adjusted for reliability and other artefacts *k* = number of samples ρ = weighted correlation coefficient; CI = confidence interval; LL = lower limit; UL = upper limit; *Q* = heterogeneity index.
*$p < .05$. **$p < .01$. ***$p < .001$.

Regarding employee voice and emotional exhaustion, stronger effect sizes were found in studies conducted in non-Western/Westernised (ρ = −0.31, 95% *CI* [−0.38; −0.23]), compared to Western/Westernised countries (ρ = −0.20, 95% *CI* [−0.31; −0.08]), as well as in studies with $N < 200$ (ρ = −0.41, 95% *CI* [−0.57; −0.22]) compared to studies with $N > 200$ (ρ = −0.24, 95% *CI* [−0.31; −0.17]). Similar effect sizes were identified for the rest of the categorical moderators examined (Table 2).

A series of meta-regression analyses were conducted where more than 10 studies were available testing the following continuous moderators: response rate, publication

Table 2. Subgroup analysis for the relationships between employee silence/employee voice outcomes and emotional exhaustion.

					95%CI	
		k	*P*	SD_ρ	[LL; UL]	*Q*
Employee silence * emotional exhaustion						
Region	Non-Western/Westernised	29	.46	.38	[.34; .56]	.54
	Western/Westernised	4	.33		[.04; .62]	
Burnout measure	MBI	30	.46	.37	[.35; .56]	1.30
	Non-MBI	3	.23		[-0.19; .58]	
>55% Female	>55% Female	7	.54	.39	[.31; .72]	1.05
	<55% Female	21	.41		[.26; .54]	
Same time point	Different time points	6	.42	.38	[.14; .64]	0.04
	Same time point	27	.45		[.32; .55]	
Quality Assessment	Fair	28	.39	.34	[.27; .49]	3.67
	Poor	8	.60		[.41; .74]	
N > 200	> 200	8	.31	.37	[.06; .53]	1.70
	< 200	25	.48		[.36; .58]	
Employee voice * emotional exhaustion						
Region	Non-Western/Westernised	17	−0.30	.18	[−0.38; −0.22]	3.51
	Western/Westernised	9	−0.16		[−0.28; −0.04]	
Burnout Measure	MBI	17	−0.24	.18	[−0.32; −0.16]	0.28
	Non-MBI	8	−0.28		[−0.39; −0.15]	
>55% Female	<55% Female	12	−0.26	.18	[−0.35; −0.15]	0.07
	>55% Female	11	−0.27		[-0.37; −0.17]	
Same time point	Different time pointS	1	−0.21	.17	[-0.52; .15]	0.06
	Same time point	24	−0.25		[−0.32; −0.19]	
Quality assessment	Fair	23	−0.27	.17	[−0.34; −0.21]	0.08
	Poor	2	−0.31		[−0.52; −0.07]	
N > 200	< 200	2	−0.39	.17	[−0.59; −0.15]	1.51
	> 200	23	−0.24		[−0.31; −0.17]	

Note: Mixed-effects analysis, *k* = number of samples, ρ = weighted correlation coefficient; CI = confidence interval; LL = lower limit; UL = upper limit. *Q* = heterogeneity index; mixed-effects analysis.

year, female percentage (continuous), sample size, mean age, mean years in the organisation, number of items in burnout measure, reliability estimates for burnout and quality assessment score (continuous). The meta-regression analyses for the relationship between employee silence and emotional exhaustion showed no significant continuous moderators (Table 3). For the relationship between employee voice and emotional exhaustion, the following moderators were significant: response rate, with higher response rate associated with larger effect sizes ($Q = 6.39$, $df = 1$, $p < .05$, $k = 17$) and publication year, whereby more recent studies were associated with larger effect sizes ($Q = 4.79$, $df = 1$, $p < .05$, $k = 25$) (Table 3).

Publication bias

Egger's regression intercept was statistically significant for the relationship between all employee silence/voice outcomes and burnout ($p < .001$, two tailed), suggesting the presence of publication bias. However, it was non-significant for the employee silence–burnout relationship and the employee voice – burnout relationship separately. We also calculated Orwin's fail-safe N, which was equal to 68 for the relationship between all employee silence/voice outcomes and burnout, 133 for the employee silence – burnout relationship and 47 for the employee voice – burnout relationship (using 0.10 as a criterion for a trivial correlation).

The standard error funnel plots for the observed studies (Supplemental material 5, Figures 2a, 3a, and 4a) indicated a degree of asymmetry. Duval and Tweedie's Trim and

Table 3. Meta-regression analyses: individual moderator model.

Moderator	*k*	Coefficient	*SE*	95%CI [LL; UL]	*Z-value*	*Q Statistics*
Employee silence and emotional exhaustion						
Response rate	23	0.006	.006	[−0.006; 0.02]	1.01	$Q = 1.01$, $df = 1$, $p = .32$
Publication year	33	−0.04	.02	[−0.08; 0.009]	−1.59	$Q = 2.52$, $df = 1$, $p = .11$
% Female participants	28	0.0003	.004	[−0.008; 0.009]	0.07	$Q = 0.00$, $df = 1$, $p = .95$
Sample size	33	−0.0001	.0003	[−0.0007; 0.0005]	−0.28	$Q = 0.08$, $df = 1$, $p = .78$
Age (mean)	21	−0.006	.009	[−0.02; 0.01]	−0.63	$Q = 0.40$, $df = 1$, $p = .53$
Years in the organization (mean)	19	−0.01	.02	[−0.04; 0.02]	−0.71	$Q = 0.51$, $df = 1$, $p = .48$
N of items in burnout measure	32	0.01	.01	[−0.01; 0.04]	0.97	$Q = 0.94$, $df = 1$, $p = .33$
Reliability coefficient for burnout	31	−0.45	.57	[−1.57; 0.67]	−0.78	$Q = 0.61$, $df = 1$, $p = .43$
Quality assessment score	33	−0.07	.04	[−0.15; 0.01]	−1.73	$Q = 2.99$, $df = 1$, $p = .08$
Employee voice and emotional exhaustion						
Response rate	17	−0.006	.002	[−0.01; −0.001]	−2.53	$Q = 6.39$, $df = 1$, $p = .01$
Publication year	25	−0.02	.009	[−0.04; −0.002]	−2.19	$Q = 4.79$, $df = 1$, $p = .03$
% Female participants	23	−0.002	.002	[−0.005; 0.0007]	−1.49	$Q = 2.21$, $df = 1$, $p = .14$
Sample size	25	−0.000	.0001	[−0.0002; 0.0001]	−0.40	$Q = 0.16$, $df = 1$, $p = .69$
Age (mean)	23	0.01	.006	[−0.0007; 0.02]	1.85	$Q = 3.44$, $df = 1$, $p = .06$
Years in the organization (mean)	15	0.01	.01	[−0.01; 0.03]	0.82	$Q = 0.68$, $df = 1$, $p = .41$
N of items in burnout measure	25	0.001	.007	[−0.01; 0.02]	0.17	$Q = 0.03$, $df = 1$, $p = .87$
Reliability coefficient for burnout	25	−0.31	.68	[−1.64; 1.02]	−0.46	$Q = 0.21$, $df = 1$, $p = .65$
Quality assessment score	25	−0.004	.03	[-0.07; 0.06]	−0.12	$Q = 0.01$, $df = 1$, $p = .91$

Fill method suggested that there were 13 studies missing on the left side of the funnel plot for all employee silence/voice outcomes, as well as 13 studies missing on the left side of the funnel plot for employee silence and 9 studies missing on the right side for employee voice separately. The adjusted stardard error funnel plots with observed and imputed studies (Supplemental material 5, Figures 2b, 3b, and 4b) showed a more symmetrical distribution of studies.

Discussion

The aim of this review and meta-analysis was to examine the correlations between employee silence/voice outcomes and job burnout. Based on diverse occupational samples found in 84 studies that reported correlations over a 21-year period, our findings indicated that overall, employee silence/voice outcomes are significantly associated to burnout, with greater silence and less voice linked to greater job burnout, in line with the results of previous meta-analyses. Examining voice and silence outcomes separately, a substantial overlap was observed between emotional exhaustion and employee silence and a small overlap was found between emotional exhaustion and employee voice. A significant moderate effect size was observed between cynicism/depersonalisation and employee silence, while the relationship between cynicism/depersonalisation and employee voice was not significant. Small effects were observed between professional efficacy with both employee silence and voice. Subgroup analysis showed that studies of poorer quality, conducted in non-Western/Westernised countries, studies using the MBI, with sample sizes greater than 200 and more than 55% female participants resulted in larger effect sizes between employee silence and emotional exhaustion compared to studies with higher quality scores, conducted in Western/Westernised countries, using non-MBI measures, with sample sizes less than 200 participants and with less than 55% female participants. In terms

of separate continuous moderators, studies with higher response rates, and more recently published studies resulted in higher effect sizes between emotional exhaustion and employee voice. These findings extended the literature in the following ways.

First, our findings support that there is a statistically significant relationship with both of the examined variables and burnout. In terms of the main meta-analysis, the overall positive effect for all employee silence/voice (ρ = .36) was closer to that between employee silence and burnout reported in the previous meta-analyses (ρ = .40; Hao et al., 2022, ρ = .32; Sherf et al., 2021) compared to the previously reported effect size between employee voice and burnout (ρ = −0.11; Sherf et al., 2021). Examining employee silence separately, the moderate effect size found between burnout and reasons for employee silence was similar to that reported by Hao et al. (2022) and slightly larger than that reported by Sherf et al. (ρ = .32), supporting the evidence that greater silence is indeed associated with greater burnout. The small-effect sizes between employee voice outcomes and burnout are slightly larger than those found by Sherf et al. (2021); however, this could be partly explained by differences in the coding of data used in this meta-analysis. Nevertheless, the findings support that the strength of the relationship between employee silence and burnout is larger compared to employee voice. Despite differences in the average strength of the effect sizes, our meta-analysis replicated a similar pattern in terms of the direction and strength of relationships on a larger pool of participants for both employee silence (N = 13,633 compared to N = 5,318 in Sherf et al., 2021 and N = 4,451 in Hao et al., 2022) and for employee voice (N = 12,361 compared to N = 5,753 in Sherf et al., 2021).

Second, our analysis showed that the relationships between the various silence/voice outcomes and components of burnout are nuanced and not straightforward. Although the available evidence largely includes studies measuring emotional exhaustion, the analysis showed that emotional exhaustion and cynicism/depersonalisation are more strongly associated with greater employee silence than employee voice, whereas reduced professional efficacy is equally related to silence and voice, albeit in the opposite directions. Effect sizes for studies measuring total burnout scores were generally larger than the effects for the three dimensions for silence (i.e. greater silence was associated with greater burnout) and voice (greater voice was associated with less burnout), though the confidence intervals for silence was narrower compared to voice, in spite of the equally limited number of studies. When multidimensional measures such as the MBI are combined to produce one overall score, the aggregated variance from the multiple combined dimensions can result in an overestimation of the effect sizes, while at the same time masking the true effect of each component (Maslach et al., 2001). Thus, the 'latent variable effect' can occur when the individual dimensions are combined, contributing to stronger relationships with both silence and voice. Also, the interaction between the different components of burnout can result in a synergy effect whereby the overall effect ends up being stronger than the sum of its parts. Thus, future research is required to investigate further whether the latent variable effect is an artefact of measurement or whether persons who for example score higher on all the different burnout components indeed experience more silence due to feeling emotionally drained, depersonalised and ineffective at the same time.

Third, our meta-analysis uniquely contributed to the understanding of factors that might be influencing the relationships between emotional exhaustion with employee silence and employee voice outcomes, highlighting that demographic and design-level decisions can influence the strength of the effect sizes. The potential influence of non-Western/Westernised countries versus Western/Westernised countries on effect sizes between emotional exhaustion with both employee silence and voice could be indicative of potential cultural effects, methodological effects or a combination of both (e.g. restraint of expression of negative emotions combined with measurement bias). Similar differences have been found between Eastern and Western countries in other meta-analyses in other areas in psychology (e.g. psychotherapy in adult depression, Tong et al., 2023) as well as in work/organisational psychology (e.g. perceived organisational support; Rockstuhl et al., 2020). Sample size was also found to influence the effect sizes, though towards opposite directions for silence and voice; where studies with sample sizes greater than 200 participants yielded larger effects for silence and EE, smaller effect sizes for voice and emotional exhaustion were observed. This could be related to the sample sizes of the individual studies for silence/voice as for example, larger sample sizes are more likely to detect smaller effects and/or regress to the mean, whereas smaller samples are more susceptible to overestimating the real effects due to biases and/or confounding factors (Button et al., 2013; Cohen, 1988). This is in line with our findings that the combined effect size for emotional exhaustion and employee silence was larger for studies rated as poor compared to studies rated as fair at the quality assessment. Stronger effects for greater silence and greater emotional in studies with > 55% female participants suggest a more pronounced impact for women. This could account for female employees feeling they are more often expected to engage in emotional labor (Hochschild, 1983) or are more afraid to be stigmatised as 'troublemakers' when speaking up at work in organisations with high levels of perceived gender inequality (Cooper et al., 2021). This finding is also in line with the evidence showing higher emotional exhaustion levels among women compared to men (Purvanova & Muros, 2010) and stronger effect sizes for depression and burnout in studies with a higher proportion of female participants (Meier & Kim, 2022).

The results of the current meta-analysis also found that the identified moderators did not account for all the heterogeneity in the observed effects. Differences in how silence and voice are operationalised across studies may explain some of the unexplained variability. For example, our review found a stronger effect size for employee silence motives compared to employee silence content, suggesting that further clarification is needed of what employee silence is and whether different measures capture different aspects of its behavioral, emotional or cognitive components. For example, asking participants whether they remain silent about specific issues does not capture exactly the same information as asking them if they remain silent due to a specific – and usually negatively framed – reason. The range of different measures used combined with the limited number of available studies did not allow any meaningful subgroup analyses to compare effect sizes by the measure used. However, the available evidence indicates that this might be one cause of heterogeneity, especially given that some questionnaires have been developed to capture specific aspects of

silence/voice in a particular occupational group (e.g. Tangirala & Ramanujam, 2008, originally for patient safety). In addition, certain aspects of employee silence motives questionnaires include explicit mentions to emotional states and experiences that are commonly associated with burnout, such as low self-efficacy, acquiescence and a need to avoid extra burden/workload (Knoll & Van Dick, 2013). Moreover, negative affectivity is implicit in most employee silence motives measures (e.g. remain silent out of fear, need to protect oneself) - but not so in employee silence content measures (e.g. Detert & Edmondson, 2011). As employee silence motives questionnaires are also capturing aspects of negative affectivity towards one's workplace, we cannot dismiss that usually, higher burnout levels coexist with more challenges in the workplace (i.e. there might be more to disagree with or feel afraid to talk about) or with more emotionally challenging professions (e.g. healthcare professionals, teachers). Similarly, meta-analytic evidence has shown that psychological safety is much more strongly associated to employee silence than to employee voice (Hao et al., 2022; Sherf et al., 2021), which could partially explain the stronger effect sizes between employee silence and burnout compared to employee voice and burnout.

The evidence examined in this review also highlights the complexities surrounding employee voice, indicating that employee voice consists of overlapping behavioural, cognitive and emotional constructs rather than being one construct in itself. Jing et al. (2014) reported a positive relationship between greater aggressive voice and greater emotional exhaustion; Duan et al. (2021) reported a positive relationship between greater self-interested voice and greater emotional exhaustion; prohibitive voice on the other hand can have a positive (e.g. Akhtar et al., 2017) and sometimes a negative relationship with burnout (e.g. Qin et al. 2014; Study 1). On the other hand, where measures of promotive voice were used (e.g. offer ideas, suggestions and solutions aimed at improving organisational processes, as in LePine & Van Dyne, 1998), negative relationships with burnout have been consistently found. This further highlights the importance of the emotional components involved in employee voice, as the behaviour alone can be indicative of positive or negative experiences, and the two might be occurring simultaneously. Thus, assuming that voice is always a positive experience would be inaccurate and further research is required.

The findings of the current review add to the increasing body of evidence suggesting that voice and silence most likely occur simultaneously at multiple levels, in concurrent situations and within/across different working relationships. To our knowledge, there is no evidence on whether individuals who score higher on the existing questionnaires for silence have in fact engaged in more silence behaviours or if, for example, due to high burnout levels, their evaluations are based on their overall experience of an organisational culture valorising silence and punishing voice. A main challenge is to understand whether employee silence motives are in fact capturing an aspect of psychological safety at work (e.g. what is the norm in the organisation/team) combined with the underlying negative affect (e.g. how likely the employee is to perceive their work environment as negative) rather the employee's individual intention to (not) engage in speaking-up behaviors.

The lack of longitudinal studies in this area means that inferences cannot be made about the direction of causality between burnout and employee silence or employee

voice. While two recent meta-analyses have positioned burnout as an outcome of silence and/or voice behaviors (Hao et al., 2022; Sherf et al., 2021), this proposition is not currently supported by evidence due to a lack of longitudinal research. The one known longitudinal study by Knoll et al. (2019) showed that when burnout preceded all four silence types the effects were stronger than when silence preceded burnout, suggesting that burnout may be an antecedent of silence. From a theoretical point of view, it can be hypothesised that burnout and silence interact in a reciprocal, dynamic bidirectional relationship that occurs in a feedback loop, which can be reinforcing or balancing (Veldhuis et al., 2020). Future longitudinal research is needed in order to establish the causal relationships between employee silence, employee voice and the components of burnout.

The evidence from this meta-analysis raises some important implications for policy and practice in organisations. The stronger relationships between greater emotional exhaustion and employee silence compared to employee voice suggest that organisations should not limit their attempts to ensuring the presence of voice channels alone or monitoring how frequently employees speak up. Where high levels of burnout are recorded, it is likely that employees do not say what they really think, but what they think is permitted or what the management wants to hear, meaning that most voice is acquiescent voice. Particularly in industries characterised by high levels of burnout – such as healthcare – organisations should be concerned that the absence of disagreement is more likely due to emotional exhaustion and cynicism rather than a belief that what the management does is right. This suggests that silence – and not voice – might be more important both as a psychological construct and as a priority for organisational management, and that the two cannot be considered two ends of a spectrum.

Limitations

We recognise that there are a number of limitations of the current systematic review and meta-analysis. First, the studies included in this meta-analysis reported on retrospective self-reported ratings of employee silence and employee voice, introducing recall bias. Although various methods for correcting for recall bias have been suggested (Bong et al., 2024; Raphael, 1987), this remains a challenge in retrospective studies. Additionally, while these measures can be viewed as proxies for silence/voice behaviours in the workplace, they should not be treated as evidence of a relationship between burnout and actual behaviors. As we excluded studies where the measure of silence/voice was not specified or could not be retrieved, unpublished data and peer-reviewed studies not in English, this might have narrowed the scope of findings and reinforced existing cultural or theoretical biases. Acknowledging these limitations is important, and future research should explore this literature to assess whether the exclusion of these studies has led to systematic biases in effect sizes, conceptual understandings of voice and silence, or the generalisability of findings across different cultural and organisational contexts. For example, future reviews can conduct a broad call for unpublished data to assess whether their exclusion systematically affects effect size estimates or theoretical interpretations.

Though the funnel plots suggested minimal bias, the file–drawer problem is still a probability with significant or positive results being more likely to be published than studies with non-significant or negative results, leading to an overestimation of the effect sizes. As papers were also excluded due to missing data, future meta-analyses could explore more advanced imputation techniques (e.g. multiple imputation, Bayesian estimation) to further mitigate data loss. Moreover, the available evidence is limited to emotional exhaustion mainly ($k = .68$) which fundamentally limits our understanding to one component of burnout and more research is needed to better understand the association with cynicism/depersonalisation and professional efficacy.

Also, voice and silence definitions in the current literature assume face-to-face interactions, with the motives and behaviors described predating the current remote-working age. Last, the initial low agreement in the quality appraisal as well as the challenges in applying certain of the assessment criteria suggests that the Quality Assessment Tool for Observational Cohort and Cross-Sectional Studies might not have been the best fit for the studies included in this review and different quality assessment tools might be better to use in future. Thus, further research is needed to understand how employee silence and voice might be different in the digital work era.

Conclusion

The evidence consistently shows a larger overlap between greater employee silence and greater emotional exhaustion and cynicism/depersonalisation, compared to employee voice, suggesting that addressing employee silence might be more beneficial for managing job burnout than focusing on employee voice. The subgroup analysis revealed larger effect sizes particularly in non-Western regions and studies using the Maslach Burnout Inventory, indicating that study-level factors need to be considered carefully. However, the available evidence is mainly cross-sectional and limited to emotional exhaustion, and future research in order to establish the causal relationships between employee silence, employee voice and the components of burnout.

Notes

1. At the PROSPERO registration stage, an additional aim to examine the direction of the relationship between burnout and employee silence outcomes and between burnout and employee voice outcomes was pre-registered. However, this was removed given the absence of longitudinal studies in the area except for one study by Knoll et al. (2019).
2. When data was entered in the CMA, the direction of the effect sizes was in the opposite direction of the conceptual coding (i.e. positive direction for employee silence outcomes and negative direction for employee voice and voice climate outcomes).
3. A main meta-analysis was conducted on the original coefficients as a form of sensitivity analysis to identify and track significant changes due to positive/negative coding of voice and silence. A summary of the main meta-analysis using the original coefficients is available in Supplemental Material 6.
4. Sensitivity analysis (one study removed) showed that the effect size became statistically significant when the Abied et al. (2019) study was removed at $\rho = .21$, 95% CI [.13, .28].
5. Sensitivity analysis (one study removed) showed that the effect size became marginally significant when the Abied et al. (2019) study was removed at $\rho = .11$, 95% CI [.01, .20].
6. Includes studies that approached burnout as unidimensional and reported effect sizes for overall burnout using the MBI ($k = 12$), the Malach-Pines (2005) burnout measure

($k = 2$), the BBI ($k = 1$), the OLBI ($k = 1$) and the Dolan et al. (2015) one-item burnout measure ($k = 1$).

7. MBI studies ρ = –.26, 95% CI [–.44, –.06], $k = 5$, $n = 2{,}954$; Non-MBI studies ρ = –.66, 95% CI [–.81, –.42], $k = 2$, $n = 399$.

Author contributions

OL, JJ and DOC carried out conceptualisation; OL, BG, PK and CM contributed to data curation; OL, JJ and DOC contributed to formal analysis; OL, BG, PK and CM contributed to investigation; OL, JJ and DOC contributed to methodology; OL carried out project administration; OL provided resources; OL, JJ and DOC supervised the study; OL carried out visualisation; OL contributed to writing – original draft; OL, JJ, CK and DOC contributed to writing – review and editing.

Disclosure statement

The authors report there are no competing interests to declare.

Funding

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

References

Abied, E. L. A. A., & Khalil, H. M. A. (2019). The impact of organizational silence on job burnout: Applied study on doctors and nursing staff at Ain Shams University Hospitals. *Journal of Contemporary Economic Studies*, *10*(1), 69–104. https://jces.journals.ekb.eg/article_50913.html

Akhtar, S., Luqman, R., Raza, F., Riaz, H., Tufail, H. S., & Shahid, J. (2017). The impact of workplace incivility on the psychological wellbeing of employees through emotional exhaustion. *European Online Journal of Natural and Social Sciences*, *6*(3), 492–507.

Alarcon, G., Eschleman, K. J., & Bowling, N. A. (2009). Relationships between personality variables and burnout: A meta-analysis. *Work & Stress*, *23*(3), 244–263. https://doi.org/10.1080/02678370903282600

Bakker, A. B., & Demerouti, E. (2007). The Job Demands-Resources model: State of the art. *Journal of Managerial Psychology*, *22*(3), 309–328. https://doi.org/10.1108/02683940710733115

Belias, M., Rovers, M. M., Reitsma, J. B., Debray, T. P. A., & IntHout, J. (2019). Statistical approaches to identify subgroups in meta-analysis of individual participant data: A simulation study. *BMC Medical Research Methodology*, *19*(1), 183. https://doi.org/10.1186/s12874-019-0817-6

Bong, S., Lee, K., & Dominici, F. (2024). Differential recall bias in estimating treatment effects in observational studies. arXiv. https://doi.org/10.48550/arXiv.2307.02331

Borenstein, M. (2022). In a meta-analysis, the I-squared statistic does not tell us how much the effect size varies. *Journal of Clinical Epidemiology*, *152*, 281–284. https://doi.org/10.1016/j.jclinepi.2022.10.003

Borenstein, M., Hedges, L. V., Higgins, J. P. T., & Rothstein, H. R. (2009). Introduction to meta-analysis. John Wiley & Sons. https://books.google.gr/books?hl=el&lr=&id=pdQnEAAAQBAJ

Button, K. S., Ioannidis, J. P. A., Mokrysz, C., Nosek, B. A., Flint, J., Robinson, E. S. J., & Munafò, M. R. (2013). Power failure: Why small sample size undermines the reliability of neuroscience. *Nature Reviews. Neuroscience*, *14*(5), 365–376. https://doi.org/10.1038/nrn3475

Chamberlin, M., Newton, D. W., & Lepine, J. A. (2017). A meta-analysis of voice and its promotive and prohibitive forms: Identification of key associations, distinctions, and future research directions. *Personnel Psychology*, *70*(1), 11–71. https://doi.org/10.1111/peps.12185

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Routledge.
Connelly, C. E., Zweig, D., Webster, J., & Trougakos, J. P. (2012). Knowledge hiding in organizations. *Journal of Organizational Behavior*, *33*(1), 64–88. https://doi.org/10.1002/job.737
Cooper, R., Mosseri, S., Vromen, A., Baird, M., Hill, E., & Probyn, E. (2021). Gender matters: A multilevel analysis of gender and voice at work. *British Journal of Management*, *32*(3), 725–743. https://doi.org/10.1111/1467-8551.12487
Detert, J. R., & Edmondson, A. C. (2011). Implicit voice theories: Taken-for-granted rules of self-censorship at work. *Academy of Management Journal*, *54*(3), 461–488. https://doi.org/10.5465/amj.2011.61967925
Dolan, E. D., Mohr, D., Lempa, M., Joos, S., Fihn, S. D., Nelson, K. M., & Helfrich, C. D. (2015). Using a single item to measure burnout in primary care staff: A psychometric evaluation. *Journal of General Internal Medicine*, *30*(5), 582–587. https://doi.org/10.1007/s11606-014-3112-6
Donaghey, J., Dundon, T., Cullinane, N., Dobbins, T., & Hickland, E. (2019). Managerial silencing of employee voice. In P. Holland, J. Teicher, & J. Donaghey (Eds.), *Employee Voice at Work* (pp. 113–128). Springer. https://doi.org/10.1007/978-981-13-2820-6_6
Duan, J., Xu, Y., Wang, X., Wu, C., & Wang, Y. (2021). Voice for oneself: Self-interested voice and its antecedents and consequences. *Journal of Occupational and Organizational Psychology*, *94*(1), 1–28. https://doi.org/10.1111/joop.12334
Duval, S., & Tweedie, R. (2000). Trim and fill: A simple funnel-plot–based method of testing and adjusting for publication bias in meta-analysis. *Biometrics*, *56*(2), 455–463. https://doi.org/10.1111/j.0006-341X.2000.00455.x
Egger, M., Davey Smith, G., Schneider, M., & Minder, C. (1997). Bias in meta-analysis detected by a simple, graphical test. *BMJ (Clinical Research ed.)*, *315*(7109), 629–634. https://doi.org/10.1136/bmj.315.7109.629
Feng, S., Shu-Xun, H., Jia-Liang, Z., Dong-Feng, R., Zheng, C., & Jia-Guang, T. (2014). Quality assessment tool for observational cohort and cross-sectional studies. *PLoS One*, *9*(11), e111695. https://doi.org/10.1371/journal.pone.0111695
Hagedoorn, M., Van Yperen, N. W., Van de Vliert, E., & Buunk, B. P. (1999). Employees' reactions to problematic events: A circumplex structure of five categories of responses, and the role of job satisfaction. *Journal of Organizational Behavior*, *20*(3), 309–321. https://doi.org/10.1002/(SICI)1099-1379(199905)20:3%3C309::AID-JOB895%3E3.0.CO;2-P
Hao, L., Zhu, H., He, Y., Duan, J., Zhao, T., & Meng, H. (2022). When is silence golden? A meta-analysis on antecedents and outcomes of employee silence. *Journal of Business and Psychology*, *37*(5), 1039–1063. https://doi.org/10.1007/s10869-021-09788-7
Higgins, J. P. T., Thomas, J., Chandler, J., Cumpston, M., Li, T., Page, M. J., & Welch, V. A. (Eds.). (2024). *Cochrane handbook for systematic reviews of interventions* (Version 6.5, updated August 2024). Cochrane. https://www.training.cochrane.org/handbook
Hirschman, A. O. (1970). *Exit, voice, and loyalty: Responses to decline in firms, organizations, and states* (Vol. 25). Harvard university press.
Hobfoll, S. E., & Freedy, J. (2017). Conservation of resources: A general stress theory applied to burnout. In *Professional burnout* (pp. 115–129). Routledge.
Hobfoll, S. E., & Shirom, A. (2000). Conservation of resources theory: Applications to stress and management in the workplace. In R. T. Golembiewski (Ed.), *Handbook of organizational behavior.* (pp. 285–312). Marcel Dekker.
Hochschild, A. R. (1983). *The managed heart: Commercialization of human feeling*. University of California press.
Jing, R., Lin Xie, J., & Ning, J. (2014). Commitment to organizational change in a Chinese context. *Journal of Managerial Psychology*, *29*(8), 1098–1114. https://doi.org/10.1108/JMP-08-2011-0042
Knoll, M., Hall, R. J., & Weigelt, O. (2019). A longitudinal study of the relationships between four differentially motivated forms of employee silence and burnout. *Journal of Occupational Health Psychology*, *24*(5), 572–589. https://doi.org/10.1037/ocp0000143
Knoll, M., & Van Dick, R. (2013). Do I hear the whistle…? A first attempt to measure four forms of employee silence and their correlates. *Journal of Business Ethics*, *113*(2), 349–362. https://doi.org/10.1007/s10551-012-1308-4

Koutsimani, P., & Montgomery, A. (2022). Burnout and cognitive functioning: Are we underestimating the role of visuospatial functions? *Frontiers in Psychiatry*, *13*, 775606. https://doi.org/10.3389/fpsyt.2022.775606

Leiter, M. P., & Maslach, C. (2016). Latent burnout profiles: A new approach to understanding the burnout experience. *Burnout Research*, *3*(4), 89–100. https://doi.org/10.1016/j.burn.2016.09.001

LePine, J. A., & Van Dyne, L. (1998). Predicting voice behavior in work groups. *Journal of Applied Psychology*, *83*(6), 853–868. https://doi.org/10.1037/0021-9010.83.6.853

Liang, J., Farh, C. I. C., & Farh, J. L. (2012). Psychological antecedents of promotive and prohibitive voice: A two-wave examination. *Academy of Management Journal*, *55*(1), 71–92. https://doi.org/10.5465/amj.2010.0176

Lipsey, M. W., & Wilson, D. B. (2001). *Practical meta-analysis*. Sage Publications, Inc.

Malach-Pines, A. (2005). The burnout measure, short version. *International Journal of Stress Management*, *12*(1), 78–88. https://doi.org/10.1037/1072-5245.12.1.78

Maslach, C., & Leiter, M. P. (2006). Burnout. In A. M. Rossi, P. L. Perrewe, & S. L. Sauter (Eds.), *Stress and quality of working life: Current perspectives in occupational health*. (Vol. 37, pp. 42–49). IAP.

Maslach, C., Schaufeli, W. B., & Leiter, M. P. (2001). Job burnout. *Annual Review of Psychology*, *52*(1), 397–422. https://doi.org/10.1146/annurev.psych.52.1.397

Meier, S. T., & Kim, S. (2022). Meta-regression analyses of relationships between burnout and depression with sampling and measurement methodological moderators. *Journal of Occupational Health Psychology*, *27*(2), 195–206. https://doi.org/10.1037/ocp0000273

Morrison, E. W. (2023). Employee voice and silence: Taking stock a decade later. *Annual Review of Organizational Psychology and Organizational Behavior*, *10*(1), 79–107. https://doi.org/10.1146/annurev-orgpsych-120920-054654

Morrison, E. W., & Milliken, F. J. (2000). Organizational silence: A barrier to change and development in a pluralistic world. *The Academy of Management Review*, *25*(4), 706–725. https://doi.org/10.5465/amr.2000.3707697

Mowbray, P. K., Wilkinson, A., & Tse, H. H. (2021). High-performance work systems and employee voice behaviour: An integrated model and research agenda. *Personnel Review*, *50*(6), 1530–1543. https://doi.org/10.1108/PR-12-2019-0692

Pinder, C. C., & Harlos, K. P. (2001). Employee silence: Quiescence and acquiescence as responses to perceived injustice. In *Research in personnel and human resources management* (pp. 331–369). Emerald Group Publishing Limited.

Purvanova, R. K., & Muros, J. P. (2010). Gender differences in burnout: A meta-analysis. *Journal of Vocational Behavior*, *77*(2), 168–185. https://doi.org/10.1016/j.jvb.2010.04.006

Qin, X., DiRenzo, M. S., Xu, M., & Duan, Y. (2014). When do emotionally exhausted employees speak up? Exploring the potential curvilinear relationship between emotional exhaustion and voice. *Journal of Organizational Behavior*, *35*(7), 1018–1041. https://doi.org/10.1002/job.1948

Raphael, K. (1987). Recall bias: A proposal for assessment and control. *International Journal of Epidemiology*, *16*(2), 167–170. https://doi.org/10.1093/ije/16.2.167

Renaud, C., & Lacroix, A. (2023). Systematic review of occupational burnout in relation to cognitive functions: Current issues and treatments. *International Journal of Stress Management*, *30*(2), 109–127. https://doi.org/10.1037/str0000279

Rockstuhl, T., Eisenberger, R., Shore, L. M., Kurtessis, J. N., Ford, M. T., Buffardi, L. C., & Mesdaghinia, S. (2020). Perceived organizational support (POS) across 54 nations: A cross-cultural meta-analysis of POS effects. *Journal of International Business Studies*, *51*(6), 933–962. https://doi.org/10.1057/s41267-020-00311-3

Scargle, J. D. (1999). Publication bias (the "file-drawer problem") in scientific inference. arXiv. https://doi.org/10.48550/arXiv.physics/9909033

Schaufeli, W. B., Desart, S., & De Witte, H. (2020). Burnout assessment tool (BAT)—Development, validity, and reliability. *International Journal of Environmental Research and Public Health*, *17*(24), 9495. https://doi.org/10.3390/ijerph17249495

Schaufeli, W. B., & Taris, T. W. (2005). The conceptualization and measurement of burnout: Common ground and worlds apart. *Work & Stress*, *19*(3), 256–262. https://doi.org/10.1080/02678370500385913

Schaufeli, W., Hakanen, J., & Shimazu, A. (2023). Burning questions in burnout research. In N. De Cuyper, E. Selenko, M. Euwema, & W. Schaufeli (Eds.), *Job insecurity, precarious employment and burnout* (pp. 127–148). Edward Elgar Publishing. https://doi.org/10.4337/9781035315888.00017

Sedgwick, P. (2013). Meta-analyses: How to read a funnel plot. *BMJ*, *346*(mar01 2), f1342–f1342. https://doi.org/10.1136/bmj.f1342

Sherf, E. N., Parke, M. R., & Isaakyan, S. (2021). Distinguishing voice and silence at work: Unique relationships with perceived impact, psychological safety, and burnout. *Academy of Management Journal*, *64*(1), 114–148. https://doi.org/10.5465/amj.2018.1428

Sonnentag, S., & Bayer, U. V. (2005). Switching off mentally: Predictors and consequences of psychological detachment from work during off-job time. *Journal of Occupational Health Psychology*, *10*(4), 393–414. https://doi.org/10.1037/1076-8998.10.4.393

Tangirala, S., & Ramanujam, R. (2008). Employee silence on critical work issues: The cross level effects of procedural justice climate. *Personnel Psychology*, *61*(1), 37–68. https://doi.org/10.1111/j.1744-6570.2008.00105.x

Tong, L., Miguel, C., Panagiotopoulou, O. M., Karyotaki, E., & Cuijpers, P. (2023). Psychotherapy for adult depression in low-and middle-income countries: An updated systematic review and meta-analysis. *Psychological Medicine*, *53*(16), 7473–7483. https://doi.org/10.1017/S0033291723002246

Van Dyne, L. V., Ang, S., & Botero, I. C. (2003). Conceptualizing employee silence and employee voice as multidimensional constructs. *Journal of Management Studies*, *40*(6), 1359–1392. https://doi.org/10.1111/1467-6486.00384

Veldhuis, G. A., Sluijs, T., van Zwieten, M. H., Bouwman, J., Wiezer, N. M., & Wortelboer, H. M. (2020). A proof-of-concept system dynamics simulation model of the development of burnout and recovery using retrospective case data. *International Journal of Environmental Research and Public Health*, *17*(16), 5964. https://doi.org/10.3390/ijerph17165964