IMPACTOS DA MODERAÇÃO DE CONTEÚDOS NA SAÚDE MENTAL DE TRABALHADORES EM PLATAFORMAS DIGITAIS: UMA REVISÃO INTEGRATIVA

Anna Laura Bastos Leal (PIBIC/FA), Matheus Viana Braz (Orientador). E-mail: mvbraz@uem.br

Universidade Estadual de Maringá, Centro de Ciências Humanas, Letras e Artes (CCH), Maringá, PR.

Área e subárea do conhecimento: Psicologia/Psicologia do Trabalho e Organizacional.

Palavras-chave: Trabalho Digital; Riscos psicossociais; Saúde do Trabalhador.

RESUMO

A moderação de conteúdo é um trabalho imprescindível para o funcionamento de plataformas digitais, sendo, muitas vezes, invisibilizado e precarizado. Este estudo teve como objetivo analisar os impactos do trabalho de moderação de conteúdos na saúde mental dos trabalhadores através de uma revisão integrativa, buscando-se entender nuances da moderação diante de marcadores de raça, classe, gênero e território. As buscas foram conduzidas no Portal de Periódicos da CAPES, na Sage Journals e no Catálogo de Teses e Dissertações da CAPES. A partir dos critérios de inclusão previamente estabelecidos, 13 artigos compuseram a amostra final, cuja análise foi organizada em torno de quatro núcleos de sentido. Chegou-se a indicativos que o sofrimento não decorre exclusivamente da exposição a conteúdos violentos ou perturbadores, mas de sua articulação com metas, pressão temporal, insuficiência de suporte, baixo reconhecimento e responsabilização individual dos moderadores. Identificou-se, também, uma assimetria geopolítica: embora países periféricos ocupem posição relevante nas cadeias globais da moderação, a literatura concentra-se no Norte Global e apresenta lacunas na caracterização de gênero, raça, classe e território. Conclui-se que a proteção da saúde mental exige intervenções sobre a organização e as relações de trabalho, não apenas estratégias individuais de adaptação ou seleção de trabalhadores considerados mais resilientes.

INTRODUÇÃO

Diariamente, plataformas digitais recebem um enorme volume de imagens, vídeos e textos publicados por seus usuários. Parte desse material contém formas de violação dos termos de uso estabelecidos pelas empresas, como discursos de ódio e violência. Para identificar e decidir o que deve permanecer, ser restringido ou removido, combinam-se sistemas automatizados com o trabalho humano de moderação de conteúdo. Os moderadores analisam, classificam e filtram essas publicações, realizando uma atividade indispensável à segurança dos ambientes

digitais e à preservação das próprias plataformas. Ainda que essencial, o trabalho de moderação é invisibilizado, terceirizado e precarizado; frequentemente, moderadores são submetidos a jornadas exaustivas e ansiogênicas, e expostos de forma contínua a conteúdos potencialmente traumáticos (Roberts, 2019).

No atual cenário de trabalho, a isenção da responsabilidade de empresas tecnológicas em relação aos seus trabalhadores (Filgueiras e Antunes, 2020) reverbera de forma grave nos indivíduos inseridos neste contexto. Muitas vezes, moderadores são empregados por plataformas de microtrabalho, denominação que reduz e deprecia o trabalho desses indivíduos: ainda que seja uma forma de trabalho fragmentada e repetitiva, há uma atividade de elaboração de sensações, emoções e produção de significados. Mesmo se esvaziada e reduzida a um funcionamento procedimental, a atividade ainda é presente e supõe um trabalho subjetivo que se inscreve no corpo e na sociedade (Braz, 2021).

A presente revisão integrativa teve como objetivo analisar os impactos do trabalho de moderação de conteúdos na saúde mental dos trabalhadores, identificando fatores de risco e protetivos, além da contemplação científica do tema.

MATERIAIS E MÉTODOS

Realizou-se uma revisão integrativa da literatura, destinada a sintetizar o conhecimento disponível e identificar lacunas sobre a saúde mental de moderadores de conteúdo (Mendes et al., 2008). Foram incluídos artigos, teses e dissertações publicados em português ou inglês entre 2020 e 2025, nos quais a saúde mental desses trabalhadores constituísse temática central. As buscas foram realizadas no Portal de Periódicos da CAPES, na Sage Journals e no Catálogo de Teses e Dissertações da CAPES, mediante combinações dos descritores “moderação de conteúdo/content moderation”, “moderador de conteúdo/content moderator”, “saúde mental/mental health”, “riscos psicossociais/psychosocial risks” e “saúde/health”. Dos 1.711 registros identificados, 1.072 permaneceram após a exclusão de duplicatas e foram avaliados pela leitura de títulos, resumos e palavras-chave. Destes, artigos que não se adequaram à temática da pesquisa e/ou não tiveram como enfoque a saúde mental de trabalhadores moderadores de conteúdo foram excluídos. Assim, 17 pesquisas foram examinadas integralmente, levando à composição da amostra final a partir de 13 artigos. A análise dos estudos foi feita de forma sistemática, permitindo a interpretação e a formulação de conclusões (Mendes et al., 2008), o que resultou no delineamento de 4 núcleos de sentido.

RESULTADOS E DISCUSSÃO

A redução significativa de 1.711 registros para 13 publicações coerentes com a temática indica que a moderação de conteúdo ainda é pouco investigada à luz de seus impactos para a saúde mental dos trabalhadores. A leitura integral dos artigos evidenciou a predominância de produções voltadas ao Norte Global, publicadas e realizadas majoritariamente em centros de pesquisa e universidades europeias. A população amostral dos artigos não seguiu um caminho diferente, sendo possível

observar dados demográficos que atestam amostras majoritariamente compostas por mulheres e pessoas brancas. As temáticas abordadas seguiam quatro principais núcleos de sentido, sendo eles: 1) os impactos psicossociais da exposição ao conteúdo; 2) o trabalho precarizado, invisibilizado e individualizado; 3) os fatores protetivos e as estratégias de enfrentamento; e, por fim, 4) as dimensões relacionais e comunitárias do trabalho de moderação.

Diferentes estudos convergem na discussão sobre trauma e seus efeitos nos moderadores de conteúdo. Este, é apontado como o alicerce da atividade de moderação, de forma que o sofrimento laboral é, na realidade, o produto comercializado neste mercado (Jones, 2025). Relatam-se, nesses estudos, sofrimento psicológico, trauma secundário, fadiga, exaustão, distúrbios do sono, pensamentos intrusivos, isolamento social e alterações nas relações familiares. O sofrimento não decorre apenas dos conteúdos moderados, mas de sua articulação com metas de produtividade, pressão temporal, ferramentas inadequadas, baixo reconhecimento e insuficiência de suporte (Spence, Martellozo e DeMarco, 2024).Há uma tendência à individualização dos riscos, o que reverbera na autopercepção dos trabalhadores e estimula a internalização da responsabilidade pelas comunidades virtuais, intensificando o sofrimento emocional diante dos conteúdos moderados.

A ideia de individualização não se limita às políticas empresariais. Há propostas interventivas que envolvem modificações nas ferramentas de moderação e a remodelação das relações dos trabalhadores, tanto entre si quanto em relação à gestão. Aspectos relacionais foram mais aprofundados em artigos que evidenciaram o trabalho de moderação voluntário. A construção de comunidades virtuais revela uma tendência à garantir a continuidade da plataforma, um senso de coletividade e a admiração pela conexão entre diferentes indivíduos. Apesar de tais padrões comportamentais serem mais realçados em estudos com trabalhadores voluntários, sua replicabilidade não deve ser ignorada na construção de estratégias envolvendo a saúde mental de moderadores comerciais de conteúdo.

Por fim, em conjunto, os achados demonstram que a saúde mental dos moderadores deve ser compreendida como uma questão estrutural e organizacional: o sofrimento é produzido pela articulação entre exposição, condições de contratação, organização do trabalho e políticas de gestão, exigindo intervenções coletivas e responsabilização das empresas.

CONCLUSÕES

A revisão evidenciou que os impactos da moderação de conteúdo ultrapassam a exposição a materiais violentos ou perturbadores, relacionando-se também à precarização, à pressão por produtividade, ao baixo reconhecimento e à individualização dos riscos. Neste contexto, a individualização e responsabilização do moderador de conteúdo se associou de forma expressiva a processos de adoecimento. A saúde mental deve, portanto, ser compreendida como uma questão estrutural e organizacional, demandando limites de exposição, pausas efetivas, apoio especializado, fortalecimento dos coletivos e responsabilização das plataformas e empresas terceirizadas.

Embora as experiências de moderação voluntária não sejam diretamente replicáveis no trabalho comercial, elas demonstram a relevância do apoio entre pares como recursos de proteção. Por fim, a escassez de pesquisas na América Latina, na África e em outras regiões periféricas evidencia a necessidade de estudos que considerem como a terceirização, desigualdades econômicas, de gênero, raça e território configuram as experiências de sofrimento e proteção nas cadeias globais da moderação de conteúdo.

AGRADECIMENTOS

Agradeço à Fundação Araucária por financiar este projeto.

REFERÊNCIAS

BRAZ, M. V. **Heteromação e microtrabalho no Brasil**. Sociologias, Porto Alegre, ano 23, n. 57, p. 134-172, 2021. Disponível em: https://doi.org/10.1590/15174522-111017. Acesso em: 26 set. 2026.

SPENCE, R.; MARTELLOZO, E.; DEMARCO, J. Content Moderator Coping Strategies: Associations With Psychological Distress, Secondary Trauma and Well-Being. **Journal of Media Psychology:** Theories, Methods, and Applications, v. 38, n. 2, p 1-11, 2024. Disponível em: https://dx.doi.org/10.1027/1864-1105/a000454. Acesso em: 26 set. 2026.

FILGUEIRAS, V.; ANTUNES, R. Plataformas digitais, Uberização do trabalho e regulação no Capitalismo contemporâneo. **Contracampo**, Niterói, v. 39, n. 1, p.27-43, 2020. Disponível em: http://dx.doi.org/10.22409/contracampo.v39i1.38901. Acesso em: 26 set. 2026.

MENDES, K. D. S.; SILVEIRA, R. C. C. P.; GALVÃO, C. M. Revisão integrativa: método de pesquisa para a incorporação de evidências na saúde e na enfermagem. **Texto Contexto Enferm.**, Florianópolis, v. 17, n. 4, p. 758-764, 2008. Disponível em: https://doi.org/10.1590/S0104-07072008000400018. Acesso em: 26 set. 2026.

ROBERTS, S. T. **Behind the screen:** content moderation in the shadows of social media. Yale University Press, 2019.

JONES, C. The "Unsung Heroes" of the "Infocalypse": Company Representations of Commercial Content Moderators. **Television & New Media**, v. 26, n. 3, p. 300-317. Disponível em: https://doi.org/10.1177/15274764241266253. Acesso em: 26 set. 2026.