Content moderator coping strategies: associations with psychological distress, secondary trauma and wellbeing

Abstract

Content moderators (CMs) apply policy set by platforms to protect users from harmful content. It is a stressful job, associated with reduced mental health and wellbeing. In this study, an anonymous survey was used to demonstrate most CMs cope by seeking support from colleagues and this is associated with lower psychological distress and secondary trauma and higher wellbeing whereas increased smoking and alcohol consumption is associated with increased symptomology. Wellbeing services were not related to a reduction in psychological distress or trauma. We argue these results fit within a framework of trauma-informed working and provide evidence for its utility in the trust and safety sector. They also highlight the need for continued research into 'what works' to support the resilience of frontline staff.

Keywords: content moderators, coping strategies, wellbeing, trauma-informed

Introduction

Content moderators (CMs) enforce rules which tell users how to act on particular online platforms, or sort through material deemed offensive or harmful, deciding on what is acceptable (Gerrard, 2022). Without them and their oversight, platforms across the internet could host harmful content, rendering them unsafe and unusable (Drootin, 2021). In 2014, Chen estimated there were over 100,000 paid CMs worldwide and it is likely this number has increased substantially since. Additionally, platforms such as Reddit and Twitch depend on volunteer CMs to keep their sites usable (Caplan, 2018).

There is growing evidence that working within content moderation is stressful and the role of CM has been associated with a number of mental health conditions (e.g., Dosono and Semaan, 2019; Roberts et al., 2023; Steiger et al., 2021). Indeed, Pinchevski (2023) argues CMs are recruited precisely due to their ability to be impacted by content, that is, their response to the content is what allows them to recognise it is disturbing. Thus, organisations need to be properly equipped to support staff dealing with harmful content, implementing a trauma-informed system from the point of recruitment, retention and progression (Champine et al., 2019). However, the ways in which people cope with stressful situations vary, and different approaches are associated with common mental health problems like depression and anxiety (Benatov et al., 2020; Fullana et al., 2020). As a greater understanding of how CMs are affected by work their work develops, organisations need to similarly develop their understanding of ‘what works’ in terms of keeping staff resilient, including relevant organisational factors, the strategies employees currently use, and to what extent these are protective so that organisations can capitalise on them.

Literature Review

Content moderation is one of the most crucial online services companies perform and without CMs, websites would become inundated with hostile and graphic material (Drootin, 2021). It is a stressful job which involves analysing flagged user-generated content and applying legal and policy standards to decide whether it should be removed (Gerrard, 2022). As part of their job moderators are likely to be exposed to potentially harmful content such as graphic violence, abuse and child sexual abuse material. This can be through text, images and/or video (Das et al., 2020). As an integral aspect of their role, content moderators are required to understand linguistic and cultural nuances, and recognise latent meanings or sub-texts in the material they are exposed to (Gillespie, 2018; Roberts, 2019). Despite the complexity of their responsibilities, CMs are subject to high accuracy targets and voluminous workloads (Dwoskin, Whalen & Cabato, 2019). Furthermore, many lack adequate workplace protections in terms of job security and stability and mental health provision (Barrett, 2020; Roberts, 2019), leaving them vulnerable to various stressors and harm.

Symptomology

People who work in roles where they are exposed to the suffering of others are known to experience a number of mental health problems such as secondary traumatic stress (STS) and burnout (e.g., Léonard et al., 2020; Ralph, 2020), as well as more common mental health conditions like depression and anxiety (e.g., Elliot-Davies, 2018; Steiger et al., 2021). Content moderators are not immune to the adverse effects of their work. Previous research indicates that they also experience high levels of psychological distress, burnout and emotional exhaustion, as well as lowered wellbeing, increased apathy and feelings of being underappreciated (Dosono and Semaan, 2019; Lo, 2018; Riedl, Masullo, & Whipple, 2020;

Schöpke-Gonzalez, Atreja, Shin, Ahmed, & Hemphill, 2022; Spence et al., 2023a; Wohn, 2019). As evidenced by Spence et al (2023b), they describe impacts such as intrusive thoughts and detachment, which are similar to symptoms observed in individuals who have experienced trauma.

Coping Strategies

Coping represents the behavioural and cognitive strategies individuals use in response to stress (e.g., Folkman & Lazarus, 1980). One way different coping styles can be distinguished is by the use of avoidant and approach strategies (Holahan & Moos, 1987; Roth & Cohen, 1986). Avoidant strategies ignore or avoid the problem, which can include suppression of emotions and is associated with behaviours such as increased use of illicit substances (Feil & Hasking, 2008). Conversely, approach strategies try to reduce or manage a problem and involves behaviours such as seeking support from others (Lazarus & Folkman, 1984). Individuals tend to use a mix of different strategies and may display quite different coping responses in different contexts, however they may show more of a reliance on certain types (Powers, Gallagher-Thompson & Kraemer, 2002). For example, research suggests that the use of some coping strategies is more stable over time than others, for instance the use of emotional support or substances is relatively stable over time, whilst strategies such as denial and behavioural disengagement are more likely to fluctuate (Nielsen & Knardahl, 2014).

Coping strategies and mental health

Roberts (2019) suggests that some CMs use maladaptive coping strategies, as evidenced through increased alcohol consumption and weight gain. There is a wealth of evidence demonstrating strong associations between avoidant coping strategies and deteriorating

psychological outcomes, such increased anxiety and depression (e.g., Benatov et al., 2020; Murayama et al., 2020; Richardson et al., 2020). Both alcohol use and smoking are associated with a higher prevalence of mental health problems (Jane-Llopis & Matytsina, 2009; Sullivan, Fiellin, & O'Connor, 2005), with some evidence suggesting both serve as risk factors for a range of disorders (Wootton et al., 2019; Taylor et al., 2014; Boden & Fergusson, 2011). Additionally, smoking and alcohol use are used as a means to cope with distress (Corbin, Farmer, & Nolen-Hoekesma, 2013; Friedman, 2020), with those who smoke relying more on negative coping strategies such as avoidance (Feil & Hasking, 2008; Siqueira et al., 2000). Hypothesis 1: Strategies that indicate more avoidant coping mechanisms will be associated with reduced wellbeing and increased psychological distress and secondary trauma.

Research also shows approach strategies are associated with reduced prevalence of mental health problems (e.g., Herman-Stabl et al., 1995; Lazarus & Folkman, 1984; Taylor, 2007) and that approach-based coping may actually increase positive affect and subjective wellbeing (Ben-Zur, 2009). Social support-seeking through talking to others is consistently found to predict increased wellbeing and reduced psychological distress (e.g., Lin et al., 1999; Taylor, 2007; Vig et al., 2020). Moreover, social support has been found to mediate the relationship between job stress and common mental health problems such as depression and anxiety (Mensah, 2021), including in stressful frontline professions such as nurses (Chen et al., 2019) and police (Hansson et al., 2016). Frontline workers and CMs both report that the benefits of support include sharing personal and work-related concerns as well as normalising emotional reactions to these (Powell et al., 2014; Spence et al., 2023c) whereas a lack of social support at work can be detrimental to workers (Mensah, 2021). Indeed, conflict with

other CMs and lack of managerial support were mentioned as reasons for volunteer content moderators to quit (Schöpke-Gonzalez et al., 2022). Hypothesis 2: Strategies suggestive of approach coping mechanisms such as seeking social support will be associated with reduced psychological distress, secondary trauma and greater wellbeing.

Help seeking and Coping

General help seeking from therapeutic services is associated with a greater use of support seeking, less use of avoidant coping strategies and a greater belief in the expertness of psychologists (Niegocki & Ægisdóttir, 2019; Johnston, Wild, Sanderson & Kent, 2022). Relatedly, avoidant coping has been found to be related to more negative perceptions of support (Johnston et al., 2022). This is possibly because those who rely on avoidant coping view themselves as self-reliant and do not want to rely on others to help with coping (Mikulincer & Shaver, 2011). However, the type of therapeutic service available to CMs can be variable both in terms of their ability to access it but also the quality of the support provided (Drootin, 2021) and the level of confidentiality (Biddle, 2019). CMs have reported they are hesitant to speak to counsellors about what they are exposed to fearing that discussing their experiences may inadvertently traumatise the counsellors themselves. This reluctance may be stronger when therapy is not facilitated by a specialist who possesses understanding of the challenges inherent to the role of CMs (Roberts, 2019; Spence et al., 2023c). Hypothesis 3: Wellbeing service use will be associated with reduced psychological distress, secondary trauma and higher wellbeing. Hypothesis 4: Concerns about confidentiality and whether the professional understands the pressure of the role will be associated with reduced wellbeing service use and increased psychological distress, secondary trauma and lower wellbeing.

From Individuals to Organisations

Increasingly, mental health is being viewed as a collective workplace responsibility rather than an individual one (Wong & Greenwood, 2023) and the idea of trauma-informed care is being applied to organisational settings (Manning, 2022). Trauma-informed workplaces embed the principles of safety, trustworthiness and transparency, peer support, collaboration and mutuality, empowerment, and cultural, historical and gender issues, into their practice and this has been found to increase psychological wellness in a variety of organisations (Headley et al., 2023; Keesler, 2020). This involves a shift to an organisational culture that openly acknowledges trauma, fosters safety and encourages employees to seek help, both through compassionate leadership and the availability of wellbeing services and initiatives (Greenwood & Anas, 2021). Critically for moderators, trauma-informed organisations are known to foster workforce resilience in dealing with challenging content and cases (Elisseou et al., 2024). Hypothesis 5: Being able to talk to managers and the availability of wellbeing services will be associated with reduced psychological distress and secondary trauma and increased wellbeing.

Currently there is no industrywide best practice in the trust and safety sector, however we believe that organisations should move towards utilising a trauma-informed approach as the gold standard. In 2015, the Technology Coalition tried to create guidelines to support moderators, but these were specific to moderators exposed to child sexual exploitation material and were non-binding (Technology Coalition, 2015). Similarly, trauma-informed approaches are often criticised for their lack of consistency, standards and ‘soft’ adaptation – meaning policies and practices are often nice to have, and lack monitoring following

implementation (Kezelman and Stavropoulos, 2012). Essentially, improvements to healthy, trauma-informed, effective ways of establishing the workplace are limited, and inconsistent both across, and within fields (Purtle, 2020). However, by exploring what works to foster resilience, organisations within the trust and safety sector can develop best practice guidance and capitalise on already existing successful individual coping strategies.

Method

Participants

Participants were CMs who took part in an anonymous online survey that collected information about demographics (e.g., sex, age, location), different coping strategies and organisational factors. The survey was started by 213 CMs and finished by 167 (78.8% completion rate). Overall, 188 (88.3%) CMs completed the Core-10, a measure of psychological distress, 167 (78.8%) participants completed the short Warwick-Edinburgh Mental Wellbeing Scale, and 174 (81.7%) participants completed the Secondary and Vicarious Trauma Scale. The sample was 49.8% female, 45.1% were based in Europe, the average age was 34.3 years (SD = 9.54) and the average duration in a CM role was 51.76 months (SD = 54.52). A full description of the sample can be found in Spence et al.'s (2023a) paper, which aimed to establish rates of psychological distress, secondary trauma, and wellbeing in the sample and explore how various work features (e.g., frequency of exposure, feedback) affected mental health. It did not however, explore the efficacy of strategies put into place by moderators and organisations to address or cope with any mental health issues.

Procedure

Convenience sampling was used to recruit participants, through both the research team's networks and by CMs who became acquainted with the research. Through these connections,

participants were identified and invited to take part. Additionally, snowball sampling was deployed, wherein CMs and their managers were encouraged to share the survey link with peers and colleagues who might be interested in participating. Furthermore, the survey was shared in social media posts targeting CM professionals. Participants completed the survey, which included a range of questions about different coping techniques as well as some of organisational factors which may facilitate these. Specifically, the survey questions participants were asked to answer included if they were able to talk to their manager if they found something difficult or distressing and if they actually talk to their manager, if they smoked cigarettes or drank alcohol and if this had increased since working in their current role. These were all scored *yes/no*. Other questions were dichotomised for the analysis, this included if they found something difficult or distressing at work did they talk to their colleagues about it, and if they talked to people outside of work about their job. These were scored on four points *No, and I don't want to/No, but I would like to/Yes, I mention it/Yes, we talk in-depth*, with negative responses becoming *no* and positive responses coded as *yes*. They were also asked if staff wellbeing services were available *yes/no/don't know*, if they used the wellbeing services *Never/Occasionally/Often*, with occasionally or often being recoded as *yes*, if they had any concerns about the confidentiality of their support services and if the person they talk to understands the pressures of the job. These items were answered *No/Somewhat/Yes* with somewhat or yes coded as *yes*. Lastly, they were also asked if their self-care had improved or worsened since working in their current role, which was scored *Got worse/Stayed the same/Got better* (see Table 1 for approach and avoidant strategies). Alongside these individual items, they also completed measures of wellbeing, secondary and vicarious trauma, and psychological distress. Participation was voluntary, unpaid and anonymous. All data was secured on password protected drives in line with Data Protection

and UK GDPR regulations. This research was approved by the university psychology department's ethics board.

Table 1 Here

Measures

Short Warwick Edinburgh Mental Wellbeing Scale (SWEMWBS; Tennant et al., 2007)

The SWEMWS is a self-report measure consisting of seven positively worded items regarding thoughts and feelings over the previous two weeks. Items are scored from 1 '*none of the time*' to 5 '*all of the time*', the scores are summed and then transformed into metric scores using the SWEMWBS conversion table. Scores range from 7 to 35 with higher scores indicating higher positive mental wellbeing. High wellbeing is scored 27.5 and above, whilst low wellbeing is scored 19.5 and below.

Clinical Outcomes in Routine Evaluation (Core-10; Barkham et al., 2013)

The Core-10 is a 10 item self-report measure assessing common presentations of psychological distress over the last week. Items are scored from 0 '*not at all*' to 4 '*most or all of the time*', and cover anxiety, depression, trauma, physical problems, functioning and risk to self. Scores are summed with higher scores indicating higher levels of general psychological distress. Severe psychological distress is scored 25 and above, whilst the non-clinical range is scored 10 and below.

Secondary and Vicarious Trauma Scale (SVTS)

The SVTS is a 17 item self-report measure assessing aspects of secondary and vicarious trauma developed for the current study based on related questionnaires. Items are scored from

1 '*strongly disagree*' to 5 '*strongly agree*', and covers thoughts (*most people are not trustworthy*), emotions (*I often feel sad*), and intrusion (*I find it difficult to separate my work and personal life*) associated with secondary and vicarious trauma. Scores are summed with higher scores indicating higher levels of secondary and vicarious trauma. The internal reliability of the SVTS was good ($\alpha = .83$).

Analysis

Descriptive statistics were used to establish the frequency of each coping strategy, as well as the organisational factors under investigation. Wellbeing, psychological distress and secondary and vicarious trauma were significantly correlated (see Spence et al., 2024) therefore multivariate general linear models (MGLM) were used. Psychological distress, wellbeing and secondary trauma were entered as the dependent variables and each different coping strategy was entered as an independent variable, i.e., talking to colleagues, talking to others, actually talking to managers, smoking cigarettes, drinking alcohol, increased smoking, increased drinking, changed self-care, and using wellbeing services. The relationship between being able to talk to managers, wellbeing services available, confidentiality concerns, wellbeing service understands pressure and the different outcomes was also explored. A separate MGLM was run for each predictor and a final model explored.

Results

Over 80% of participants reported talking to their colleagues, and two thirds spoke to people outside of work about their job. Almost 90% felt able to talk to management, although only 76% actually spoke to management about things that were distressing them at work. A quarter of participants smoked, and almost three quarters drank alcohol. Of the smokers, 40%

reported their smoking had increased since beginning their role and a fifth reported drinking more. Just under a quarter of participants reported their self-care had become worse since starting their role. The majority of participants had wellbeing services available through their employment, and two thirds used the wellbeing services. A fifth of participants had concerns about the confidentiality of their wellbeing services and just under a fifth thought the services did not understand the pressure they experienced as part of their role (see Table 2).

Table 2 Here

Talking to colleagues was significantly associated with less distress ($F = 22.05$, $p < .001$) and trauma ($F = 19.63$, $p < .001$) and greater wellbeing ($F = 24.67$, $p < .001$). However, talking to people outside of work was not associated with mental health (psychological distress: $F = .22$, *n.s.*; Secondary trauma: $F = .12$, *n.s.*) or wellbeing ($F = 1.23$, *n.s.*). Being able to talk to management was associated with lower psychological distress ($F = 8.06$, $p < .01$) and trauma ($F = 11.57$, $p < .001$) and greater wellbeing ($F = 8.29$, $p < .01$), despite actually talking to management not being associated with any change in psychological distress ($F = 3.32$, *n.s.*). Drinking alcohol and smoking cigarettes were not associated with psychological distress ($F = 2.55$, *n.s.*; $F = .48$, *n.s.* respectively), secondary trauma ($F = .01$, *n.s.;* $F = 2.55$, *n.s.* respectively) or wellbeing ($F = .05$, *n.s.;* $F = 1.52$, *n.s.*, respectively), however having increased alcohol intake or smoking since beginning in the role was associated with lower wellbeing ($F = 18.20$, $p < .01$; $F = 9.67$, $p < .05$ respectively) and increased psychological distress ($F = 21.99$, $p < .01$; $F = 11.69$, $p < .05$ respectively) and trauma ($F = 32.91$, $p < .01$; $F = 16.81$, $p < .01$ respectively). This was also the case for those whose self-care had worsened (psychological distress: $F = 15.71$, $p < .01$; secondary trauma: $F = 18.76$, $p < .01$, wellbeing: $F = 14.20$, $p < .01$). Having wellbeing services available through work was associated with lower psychological distress ($F = 6.33$, $p < .05$) and secondary trauma ($F = 14.30$, $p < .001$)

and greater wellbeing (F = 5.90, $p < .05$), despite the fact that using the wellbeing services was not significantly associated with anything (psychological distress: F = .34, *n.s.*; secondary trauma: F = .01, *n.s.*; wellbeing: F = .01, *n.s.*). Those who had concerns about confidentiality of the services or did not think they understood the pressure of the role had higher psychological distress (F = 5.59, $p < .05$; F = 5.10, $p < .05$, respectively) and trauma (F = 8.01, $p < .01$; F = 7.79, $p < .01$, respectively) and lower wellbeing (F = 5.95, $p < .05$; F = 7.54, $p < .01$, respectively) (see Table 3). A full factorial model of all the predictors was assessed but showed no significant interactions.

Table 3 here

A binary logistic regression found that concerns about confidentiality (B = .156, *n.s.*) or the wellbeing service understanding the pressure of the role (B = .24, *n.s.*) did not significantly predict whether someone used wellbeing services.

Discussion

The results demonstrated that behaviours indicative of avoidant and approach coping strategies were endorsed. The majority of CMs appeared to be using approach-based coping strategies, where they were talking to colleagues and management, as well as people outside their job. Talking to colleagues appeared to be the main source of support (83%), followed by management (76%) then people outside of work (69%). In part, this may reflect the use of non-disclosure agreements, although participants were not asked about these in the current study, they are often in use in CM workplaces and prevent CMs from being able to talk to others, even if they want to (Petricca, 2020).

We hypothesised that behaviours suggestive of avoidant coping such as smoking, alcohol use and reduced self-care would be associated with greater psychological distress, secondary trauma and lower wellbeing. This tended to be the case, although drinking alcohol and smoking cigarettes were not associated with mental health or wellbeing, an increase in smoking or alcohol intake since beginning their role was associated with lower wellbeing and increased psychological distress and trauma. This was also the case for those who reported worsening self-care. These behaviours were in the minority but could indicate individuals who are reliant on more avoidant based strategies (Feil & Hasking, 2008), although equally poorer mental health could be leading to these behaviours (Aldwin & Revenson,1987; Jane-Llopis & Matytsina, 2009). Regardless, these behaviours can be used by workers and employers alike to signal potential underlying distress.

We also hypothesised that strategies that indicate more adaptive coping mechanisms like seeking social support would be associated with reduced psychological distress and trauma and increased wellbeing. To some extent this was borne out, talking to colleagues was significantly associated with less distress and secondary trauma and greater wellbeing. Conversely, talking to people outside of work was not associated with either mental health or wellbeing. We also found that being able to talk to management was associated with lower psychological distress and trauma and greater wellbeing, despite *actually* talking to management not being associated with any change in psychological distress. Previous work with CMs and frontline workers in high stress jobs such as the emergency services, has found that employees in stressful roles rely on supportive colleagues because they understand what the role involves, in a way external people do not (Billings et al., 2021; Steiger et al., 2021). It is possible talking to friends and family did not result in similar benefits because CMs do

not open up about the job in the same way. Previous work has found CMs believe others do not want to know about the job or need to be protected from hearing about the realities involved (Spence et al., 2023c). Similarly, there is some evidence that leadership do not necessarily fully understand the role of content moderators (Roberts et al., 2023) and so this, in part may explain why talking to management was not associated with a reduction in psychological distress. Although talking to colleagues may have positive effects in the short-term, this sort of coping through support seeking is often seen as maladaptive in the long-term because they only change negative feelings about the situation and not the situation itself (Aldwin & Revenson, 1987). Therefore, it would be interesting to see if the effectiveness of coping strategies changes over time and if those who stay in the role long-term use different forms of coping.

Our third hypothesis was wellbeing service use will be associated with reduced psychological distress, secondary trauma and higher wellbeing. However, using the wellbeing services was not significantly associated with mental health or wellbeing. Similarly, hypothesis four was only partly correct, in that wellbeing service use was not predicted by concerns about confidentiality or the belief that they did not understand the pressure of the role, although these was associated with higher psychological distress and trauma and lower wellbeing. There are several possibilities for this finding, their negative perceptions may suggest a reliance on avoidant coping (Johnston et al., 2022), alternatively it may suggest something about their level of engagement with (rather than use of) the available services, or it may reflect more generalised concerns regarding the company and their place within it. Regardless, it is worrying that use of the wellbeing services was not associated with mental health and wellbeing. This is possibly because the type of service available was not

appropriate to the needs of CMs in order to be effective. Some CM workplaces provide generic or non-clinical wellbeing services (Drootin, 2021), and in 2020 Meta settled a lawsuit due to providing inadequate mental health support for their CMs (Dwoskin, 2020). Certainly, CMs have expressed the need for specialist support in order for it to be useful (Spence et al., 2023c). This demonstrates the need for any wellbeing initiatives to be properly implemented and evaluated.

Lastly, we hypothesised being able to talk to managers and the availability of wellbeing services will be associated with reduced psychological distress and secondary trauma and increased wellbeing. This was correct and feeling able to talk to management and having wellbeing services available were associated with improved mental health. Taken together, the results provide encouraging evidence for implementing trauma-informed working. This includes the principles of collaboration and mutuality, where compassionate leadership, embedding trauma-awareness and normalising empathetic and reflective practices is critical for staff across an organisation not just on the frontline (Huang et al., 2014). Additionally, to some extent the availability of wellbeing services may reflect organisational cultures that are more therapeutic and wellbeing focused in nature and thus, the availability of services may help promote a sense of organisational safety, which is another feature of trauma-informed workplaces. Furthermore, the majority of those surveyed spoke to colleagues when faced with distressing content at work and this was linked with better mental health and wellbeing. Peer support is one of the principles of a trauma-informed workplace and ensures a sense of connection and uniformity with colleagues; acknowledging shared experience of trauma within the workplace (Huang et al., 2014).

Theoretical Implications

In this paper we have demonstrated some of the ways in which content moderators currently cope with their work, some organisational factors that facilitate this coping and argued that adopting a trauma-informed approach would be beneficial for the trust and safety sector. Certainly, the factors associated with improved psychological functioning fit within a trauma-informed framework. The results suggest workplaces should cultivate trusted relationships between employees, and foster collaboration and teamwork (Office of the Surgeon General, 2022), this should not only be between frontline workers but also with management. Compassionate leadership has been highlighted as an important part of trauma-informed workplaces and has been positively associated with employee well-being, positive views of the organisation as a good workplace, and willingness of employees to work effectively for the best outcomes of the organisation (Alonso, 2022). Workplace factors that are associated with employee coping should be measured. Only a minority reported increased drinking and/or smoking and reduced self-care but these could be used as markers for increased distress. Indeed, both active and avoidant coping behaviours may be important behavioural indicators of resilience and wellbeing organisations can use. The results also demonstrate an ongoing need for continued research into, and refinement of, the wellbeing services offered to CMs alongside a commitment to evaluating and monitoring the effects of changes following implementation (Fernandez et al., 2023; Kezelman and Stavropoulos, 2012). Part of the difficulty ensuring high quality and effective trauma-informed workplaces lies with different implementation methods, theoretical underpinnings, knowledge base in the work force, and monitoring plans. Organisations purport that they are trauma-informed without a standard, consistent level of trauma-informed, or trauma- considered policies and practice (Fernandez et al., 2023). Regularly, asking for employee feedback and enacting real, meaningful change

which is then properly evaluated could contribute to employee feelings of empowerment, another principle of trauma-informed workplaces, and the development of best practice guidelines, which are currently lacking.

Limitations and Future Directions

This study is limited in that it did not directly ask about coping styles, but instead asked about behavioural indicators that may signify these. Nevertheless, relying on observable behaviour as an indicator of coping and possible functioning may be more useful for organisations who are trying to ascertain levels of wellbeing among their staff. The average duration in role was longer than expected (approximately 51 months), 52.8% of moderators had been in their role for less than two years, but equally 47.2% were in role for longer. This suggests there is a cohort that stay in post for a reasonably long time period, duration was not associated with any of the coping strategies in this study, but the correlates of working as a CM longer-term warrants further investigation. Similarly, the survey did not directly ask about trauma-informed practices at work and it was not possible to ascertain why the use of wellbeing services was not associated with improved mental health and wellbeing from the information provided. More detailed information about the type, frequency and specialism of services is needed to tease apart what type of wellbeing service best alleviates work stress in this population. In addition, publicly available research which specifically directly evaluates how trauma-informed work practices influences the mental health of trust and safety employees could help as a first step towards developing best practices across the sector.

Research Transparency Statement

The authors are not willing to share their data, but will share their analytics methods, and study materials with other researchers. The material will be available upon request.

References

Aldwin, C. M., & Revenson, T. A. (1987). Does coping help? A re-examination of the relation between coping and mental health. *Journal of Personality and Social Psychology, 53*, 337–348.

Alonso, N. (2022). *Trauma-informed compassionate leader behaviors: Showing compassion in the workplace to those that have suffered trauma* (Doctoral dissertation). https://uh-ir.tdl.org/bitstream/10657/14293/1/ALONSO-DISSERTATION-2022.pdf

Bardwell, W. A., Ancoli-Israel, S., & Dimsdale, J. E. (2001). Types of coping strategies are associated with increased depressive symptoms in patients with obstructive sleep apnea. *Sleep*, *24*, 905–909.

Barkham, M., Bewick, B., Mullin, T., Gilbody, S., Connell, J., Cahill, J., Mellor-Clark, J., Richards, D., Unsworth, G., & Evans, C. (2013). The CORE-10: A short measure of psychological distress for routine use in the psychological therapies. *Counselling and Psychotherapy Research, 13*(1), 3-13. https://doi.org/10.1080/14733145.2012.729069

Barrett, P. M. (2020). *Who moderates the social media giants? A call to end outsourcing.* New York University Stern Center for Business and Human Rights.

Benatov, J., Klomek, A. B., Shira, B., Apter, A., Carli, V., Wasserman, C., Hoven, C. W., Sarchiapone, M., Balazs, J., Bobes, J., Brunner, R., Corcoran, P., Cosman, D., Haring, C., Kahn, J., Keeley, H., Kereszteny, A., Podlogar, T., Postuvan, V., . . . Wasserman, D. (2020). Doing nothing is sometimes worse: Comparing avoidant versus approach coping strategies with peer victimization and their association to depression and suicide

ideation. *Journal of School Violence, 19*(4), 456-469. https://doi.org/10.1080/15388220.2020.1738941

Ben-Zur, H. (2009). Coping styles and affect. *International Journal of Stress Management, 16*(2), 87–101. https://doi.org/10.1037/a0015731

Biddle, S. (2019, August 16). Trauma counselors were pressured to divulge confidential information about Facebook moderators, internal letter claims. *The Intercept.* *https://theintercept.com/2019/08/16/facebook-moderators-mental-health-accenture/*

Billings, J., Seif, N. A., Hegarty, S., Ondruskova, T., Soulios, E., Bloomfield, M., & Greene, T. (2021). What support do frontline workers want? A qualitative study of health and social care workers' experiences and views of psychosocial support during the COVID-19 pandemic. *PloS ONE, 16*(9), e0256454. https://doi.org/10.1371/journal.pone.0256454

Caplan, R. (2018). Content or context moderation? Artisanal, community-reliant, and industrial approaches. *Data & Society Research Institute.* https://datasociety.net/library/content-or-context-moderation/

Chen, J., Li, J., Cao, B., Wang, F., Luo, L., & Xu, J. (2019). Mediating effects of self-efficacy, coping, burnout, and social support between job stress and mental health among young Chinese nurses. *Journal of Advanced Nursing, 76*(1), 163-173. https://doi.org/10.1111/jan.14208

Corbin, W. R., Farmer, N. M., & Nolen-Hoekesma, S. (2013). Relations among stress, coping strategies, coping motives, alcohol consumption and related problems: A mediated moderation model. *Addictive Behaviors, 38*(4), 1912-1919. https://doi.org/10.1016/j.addbeh.2012.12.005

Das, A., Dang, B., & Lease, M. (2020). Fast, accurate, and healthier: Interactive blurring helps moderators reduce exposure to harmful content. *Proceedings of the AAAI Conference on Human Computation and Crowdsourcing*, *8*(1), 33-42. https://doi.org/10.1609/hcomp.v8i1.7461

Dosono, B., & Semaan, B. (2019). Moderation practices as emotional labor in sustaining online communities: The case of AAPI identity work on Reddit. In *CHI '19: Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems* (pp. 1–13). ACM. https://doi.org/10.1145/3290605.3300372

Drootin, A. (2021). "Community guidelines": The legal implications of workplace conditions for internet content moderators. *Fordham Law Review, 90*(3), 1197-1244.

Dwoskin, E., Whalen, J., & Cabato, R. (2019, July 25). Content moderators at YouTube, Facebook and Twitter see the worst of the web – and suffer silently. *The Washington Post*. https://www.washingtonpost.com/technology/2019/07/25/social-media-companies-are-outsourcing-their-dirty-work-philippines-generation-workers-is-paying-price/

Elliott-Davies, M. (2018). *PFEW demand, capacity and welfare survey 2018*. Research and Policy Support Report (R101/2018). https://www.polfed.org/media/14060/demandcapacityandwelfaresurveyheadlinestatistics2018-06-02-19-v1.pdf

Elisseou, S., Shamaskin-Garroway, A., Kopstick, A. J., Potter, J., Weil, A., Gundacker, C., & Moreland-Capuia, A. (2024). Leading organizations - from burnout to trauma-informed resilience: A vital paradigm shift. *The Permanente Journal, 28*(1), 198-205. https://doi.org/10.7812/TPP/23.110

Feil, J., & Hasking, P. (2008). The relationship between personality, coping strategies and alcohol use. *Addiction Research & Theory, 16*(5), 526-537.

Fernández, V., Gausereide-Corral, M., Valiente, C., & Sánchez-Iglesias, I. (2023). Effectiveness of trauma-informed care interventions at the organizational level: A systematic review. *Psychological Services, 20*(4), 849–862.

Folkman, S., & Lazarus, R. S. (1980). An analysis of coping in a middle-aged community sample. *Journal of Health and Social Behavior*, *21*, 219–239.

Fullana, M. A., Hidalgo-Mazzei, D., Vieta, E., & Radua, J. (2020). Coping behaviors associated with decreased anxiety and depressive symptoms during the COVID-19 pandemic and lockdown. *Journal of Affective Disorders, 275*, 80-81. https://doi.org/10.1016/j.jad.2020.06.027

Gerrard, Y. (2022). Social media moderation: The best-kept secret in Tech. In D. Rosen (Ed.), *The social media debate* (pp. 77–95). Routledge.

Greenwood, K., & Anas, J. (2021, October 4). It's a new era for mental health at work. *Harvard Business Review*. *https://hbr.org/2021/10/its-a-new-era-for-mental-health-at-work*

Hansson, J., Hurtig, A. K., Lauritz, L. E., & Padyab, M. (2017). Swedish police officers' job strain, work-related social support and general mental health. *Journal of Police and Criminal Psychology, 32*, 128-137. https://doi.org/10.1007/s11896-016-9202-0

Headley, A. M., Witkowski, K., Remington, C., Emel Ganapati, N., & Contreras, S. L. (2023). Trauma-informed organizational climate and its impact on first responder burnout during COVID-19. *Public Administration Review.* https://doi.org/10.1111/puar.13764

Herman-Stabl, M. A., Stemmler, M., & Petersen, A. C. (1995). Approach and avoidant coping: Implications for adolescent mental health. *Journal of Youth and Adolescence, 24*, 649–665. https://doi.org/10.1007/BF01536949

Holahan, C. J. & Moos, R. H. (1986). Personal and contextual determinants of coping strategies. *Journal of Personality and Social Psychology*, *52*(5), 946-955.

Huang, L. N., Flatow, R., Biggs, T., Afayee, S., Smith, K., Clark, T., & Blake, M. (2014). *SAMHSA's Concept of Trauma and Guidance for a Trauma-Informed Approach*. *S*ubstance Abuse and Mental Health Services Administration. https://ncsacw.acf.hhs.gov/userfiles/files/SAMHSA_Trauma.pdf

Jane-Llopis, E., & Matytsina, I. (2009). Mental health and alcohol, drugs and tobacco: A review of the comorbidity between mental disorders and the use of alcohol and illicit drugs. *Drug and Alcohol Review, 25*(6), 515-536. https://doi.org/10.1080/09595230600944461

Johnston, S., Wild, J., Sanderson, K., & Kent, B. (2022). Perceptions and experiences of mental health support for ambulance trust employees. *Journal of Paramedic Practice, 14*(7), 287–296. https://doi.org/10.12968/jpar.2022.14.7.287

Keesler, J. M. (2020). Promoting satisfaction and reducing fatigue: Understanding the impact of trauma-informed organizational culture on psychological wellness among Direct Service Providers. *Journal of Applied Research in Intellectual Disabilities, 33*(5), 939-949. https://doi.org/10.1111/jar.12715

Kezelman, C., & Stavropoulos, P. (2012). *Practice guidelines for treatment of complex trauma and trauma informed care and service delivery*. Adults Surviving Child Abuse. https://ymcamanchester.org.uk/wp-content/uploads/trauma-informed-service.pdf

Lazarus, R. S., & Folkman, S. (1984). *Stress, Appraisal, and Coping*. Springer Publishing.

Léonard, M. J., Saumier, D., & Brunet, A. (2020). When the lawyer becomes traumatized: A scoping review. *SAGE Open*, *10*(3), 1–10. https://doi.org/10.1177/2158244020957032

Lin, N., Ye, X., & Ensel, W. M. (1999). Social support and depressed mood: a structural analysis. *Journal of Health and Social Behavior, 40*(4), 344-359.

Lo, C. (2018). *When all you have is a banhammer: The social and communicative work of volunteer moderators* (Doctoral thesis, Massachusetts Institute of Technology). https://cmsw.mit.edu/wp/wp-content/uploads/2018/05/Claudia-Lo-When-All-You-Have-Is-a-Banhammer.pdf

Manning, K. (2022, March 31). *We need trauma-informed workplaces*. Harvard Business Review. https://hbr.org/2022/03/we-need-trauma-informed-workplaces

Mensah, A. (2021). Job stress and mental well-being among working men and women in Europe: The mediating role of social support. *International Journal of Environmental Research in Public Health*, *18*, 2494. https://doi.org/10.3390/ijerph18052494

Mikulincer, M., & Shaver, P. (2011). An attachment perspective on interpersonal and intergroup conflict. In: J. Forgas, A. Kruglanski, & K. Williams (Eds), *The psychology of social conflict and aggression* (pp. 19-35)**.** Psychology Press. https://doi.org/10.4324/9780203803813

Murayama, Y., Yamazaki, S., Yamaguchi, J., Hasebe, M., & Fujiwara, Y. (2020). Chronic stressors, stress coping and depressive tendencies among older adults. *Geriatrics & Gerontology International*, *20*(4), 297–303. https://doi.org/10.1111/ggi.13870

Niegocki, K. L., & Ægisdóttir, S. (2019). College students' coping and psychological help-seeking attitudes and intentions. *Journal of Mental Health Counseling, 41*(2), 144-157. https://doi.org/10.17744/mehc.41.2.04

Nielsen, M. B., & Knardahl, S. (2014). Coping strategies: A prospective study of patterns, stability and relationships with psychological distress. *Scandinavian Journal of Psychology, 55*(2), 142-150. https://doi.org/10.1111/sjop.12103

Office of the Surgeon General. (2022). *The US surgeon general's framework for workplace mental health & well-being.* US Department of Health & Human Services.

Petricca, P. (2020). Commercial content moderation: An opaque maze for freedom of expression and customers' opinions. *Rivista Internazionale di Filosofia e Psicologia, 11*(3), 307-326. https://doi.org/10.4453/rifp.2020.0021

Pinchevski, A. (2023). Social media's canaries: content moderators between digital labor and mediated trauma. *Media, Culture & Society*, *45*(1), 212-221. https://doi.org/10.1177/01634437221122226

Powers, D. V., Gallagher-Thompson, D., & Kraemer, H. C. (2002). Coping and depression in Alzheimer's caregivers: Longitudinal evidence of stability. *Journals of Gerontology*: *Series B, Psychological Sciences and Social Sciences*, *57*, 205–211.

Purtle, J. (2020). Systematic review of evaluations of trauma-informed organizational interventions that include staff trainings. *Trauma, Violence & Abuse, 21*(4), 725-740. https://doi.org/10.1177/1524838018791304

Ralph, N. (2020, August 21). *Dealing with the personal impact of crimes against children*. College of Policing. https://www.college.police.uk/article/dealing-personal-impact-crimes-against-children

Richardson, C. E., Magson, N. R., Fardouly, J., Oar, E. L., Forbes, M. K., Johnco, C. J., & Rapee, R. M. (2020). Longitudinal Associations between Coping Strategies and Psychopathology in Pre-adolescence. *Journal of Youth and Adolescence, 50*(6), 1189-1204. https://doi.org/10.1007/s10964-020-01330-x

Riedl, M. J., Masullo, G. M., & Whipple, K. N. (2020). The downsides of digital labor: Exploring the toll incivility takes on online comment moderators. *Computers in Human Behavior, 107,* Article 106262. https://doi.org/10.1016/j.chb.2020.106262

Roberts, S. T. (2019). *Behind the screen: Content moderation in the shadows of social media*. Yale University Press. https://doi.org/10.2307/j.ctvhrcz0v

Roberts, S. T., Wood, S. E., & Eadon, Y. (2023, January). "We care about the internet; we care about everything": Understanding social media content moderators' mental models and support needs. In *Proceedings of the 56th Hawaii international conference on system sciences*.

Roth, S., & Cohen, L. J. (1986). Approach, avoidance, and coping with stress. *American Psychologist, 41*(7), 813–819. https://doi.org/10.1037/0003-066X.41.7.813

Schöpke-Gonzalez, A. M., Atreja, S., Shin, H. N., Ahmed, N., & Hemphill, L. (2022). Why do volunteer content moderators quit? Burnout, conflict and harmful behaviors. *New Media & Society*, *0*(0). https://doi.org/10.1177/14614448221138529

Siqueira, L., Diab, M., Bodian, C., & Rolnitzky, L. (2000). Adolescents becoming smokers: The roles of stress and coping methods. *Journal of Adolescent Health, 27*(6), 399-408. https://doi.org/10.1016/S1054-139X(00)00167-1

Spence, R., Bifulco, A., Bradbury, P., Martellozzo, E., & DeMarco, J. (2023a). Content Moderator Mental Health, Secondary Trauma, and Well-being: A Cross-Sectional Study. *Cyberpsychology, Behavior, and Social Networking, 27*(2). https://doi.org/10.1089/cyber.2023.0298

Spence, R., Bifulco, A., Bradbury, P., Martellozzo, E., & DeMarco, J. (2023b). The psychological impacts of content moderation on content moderators: A qualitative

study. *Cyberpsychology: Journal of Psychosocial Research on Cyberspace*, *17*(4). https://doi.org/10.5817/cp2023-4-8

Spence, R., Harrison, A., Bradbury, P., Bleakley, P., Martellozzo, E., & DeMarco, J. (2023c). Content moderators' strategies for coping with the stress of moderating content online. *Journal of Online Trust & Safety, 1*(5). https://doi.org/10.54501/jots.v1i5.91

Steiger, M., Bharucha, T. J., Venkatagiri, S., Riedl, M. J., & Lease, M. (2021). The psychological well-being of content moderators: The emotional labor of commercial moderation and avenues for improving support. In *CHI '21: Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems* (pp. 1–14). ACM. https://doi.org/10.1145/3411764.3445092

Sullivan, L. E., Fiellin, D. A., & O'Connor, P. G. (2005). The prevalence and impact of alcohol problems in major depression: A systematic review. *The American Journal of Medicine, 118*(4), 330-341. https://doi.org/10.1016/j.amjmed.2005.01.007

Taylor, S. E. (2007). Social support. In H.S. Friedman & R.C. Silver (Eds.), *Foundations of Health Psychology* (pp. 145–71). Oxford University Press.

Taylor, G., McNeill, A., Girling, A., Farley, A., Lindson-Hawley, N., & Aveyard, P. (2014). Change in mental health after smoking cessation: Systematic review and meta-analysis. *BMJ, 348*, g1151. https://doi.org/10.1136/bmj.g1151

Technology Coalition. (2015). *Employee resilience guidebook for handling child sexual abuse images*. Technology Coalition. https://cdn.icmec.org/wp-content/uploads/2023/04/TechnologyCoalitionEmployeeResilienceGuidebookV2January2015.pdf

Tennant, R., Hiller, L., Fishwick, R., Platt, S., Joseph, S., Weich, S., Parkinson, J., Secker, J., & Stewart-Brown, S. (2007). The Warwick-Edinburgh Mental Well-being Scale

(WEMWBS): development and UK validation. *Health and Quality of Life Outcomes, 5*(1). https://doi.org/10.1186/1477-7525-5-63

Vig, K. D., Mason, J. E., Carleton, R. N., Asmundson, G. J. G., Anderson, G. S., & Groll, D. (2020). Mental health and social support among public safety personnel, *Occupational Medicine*, *70*(6), 427–433. https://doi.org/10.1093/occmed/kqaa129

Wohn, D. Y. (2019). Volunteer moderators in Twitch micro communities: How they get involved, the roles they play, and the emotional labor they experience. In *CHI '19: Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems* (pp. 1–13). ACM. https://doi.org/10.1145/3290605.3300390

Wootton, R. E., Richmond, R. C., Stuijfzand, B. G., Lawn, R. B., Sallis, H. M., Taylor, G. M. J., Hemani, G., Jones, H. J., Zammit, S., Smith, G. D., & Munafò, M. R. (2020). Evidence for causal effects of lifetime smoking on risk for depression and schizophrenia: A Mendelian randomisation study. *Psychological Medicine*, 50(14), 2435-2443. https://doi.org/10.1017/s0033291719002678

Wong, B., & Greenwood, K. (2023, October 10). The future of mental health at work is safety, community and a healthy organizational culture. *Harvard Business Review.* https://hbr.org/2023/10/the-future-of-mental-health-at-work-is-safety-community-and-a-healthy-organizational-culture

Table 1: Categorisation of approach and avoidant strategies and organisational factors explored

Avoidant	**Approach**	**Organisational factors**
Smoking (incl. if this has increased)	Talking to colleagues	Being able to talk to management
Drinking alcohol (incl. if this has increased)	Talking to others outside work	Availability of wellbeing services
Reduced self-care	Talking to management	No concerns about confidentiality
	Using wellbeing services	Wellbeing services understand the pressure

Table 2: Frequency of content moderators endorsing different behaviours

Variable	Frequency (%)
Talk to colleagues	
Yes	121 (83.4)
No	24 (16.6)
Talk to others outside work	
Yes	99 (68.8)
No	45 (31.3)
Able to talk to management	
Yes	126 (88.1)
No	17 (11.9)
Do talk to management	
Yes	109 (76.2)
No	34 (23.8)
Smoke	
Yes	37 (24.7)
No	113 (75.3)
Smoking has increased	
Yes	15 (40.5)
No	22 (59.5)
Drink alcohol	
Yes	110 (72.8)
No	41 (27.2)
Drinking has increased	

Yes	23 (21.1)
No	86 (78.9)
Self-Care	
Same	69 (46.0)
Better	46 (30.7)
Worse	35 (23.3)
Wellbeing services available	
Yes	117 (81.3)
No	16 (11.1)
Don’t know	11 (7.6)
Use wellbeing services	
Yes	77 (66.4)
No	39 (33.6)
Concerns about confidentiality	
Yes	29 (20.3)
No	114 (79.7)
Understand the pressure	
Yes	118 (82.5)
No	25 (17.5)

Table 3: Predictors of psychological distress, wellbeing and secondary trauma

		Psychological Distress		Wellbeing		Secondary Trauma	
		Mean (se)	F	Mean (se)	F	Mean (se)	F
Colleagues	Yes	12.92 (.50)	22.05***	22.77 (.39)	24.67***	44.80 (.97)	19.63***
	No	17.06 (1.37)		19.49 (1.07)		54.50 (2.68)	
Others	Yes	13.25 (.57)	.22	22.66 (.44)	1.23	46.12 (1.12)	.12
	No	13.75 (.89)		21.74 (.69)		45.40 (1.77)	
Able Talk	Yes	12.92 (.50)	8.06**	22.77 (.39)	8.29**	44.80 (.97)	11.57***
Management	No	17.06 (1.37)		19.49 (1.07)		54.50 (2.68)	
Do Talk	Yes	12.91 (.55)	3.32	23.05 (.42)	10.34**	44.82 (1.08)	4.24*
Management	No	14.91 (.96)		20.35 (.73)		49.29 (1.89)	
Drink Alcohol	Yes	12.93 (.56)	2.55	22.34 (.44)	.05	45.86 (1.11)	.01
	No	14.63 (.91)		22.53 (.72)		46.05 (1.81)	
Smoke	Yes	14.00 (1.00)	.48	21.55 (.78)	1.52	48.66 (1.96)	2.55
	No	13.22 (.55)		22.64 (.42)		45.09 (1.07)	
Drinking has	Yes	17.85 (1.17)	21.99**	18.87 (.91)	18.20**	57.10 (2.19)	32.91**
increased	No	11.72 (.58)		23.20 (.45)		43.09 (1.09)	
Smoking has	Yes	17.47 (1.39)	11.69*	19.86 (.75)	9.67*	55.07 (2.15)	16.81**
increased	No	10.94 (1.31)		23.05 (.70)		43.00 (2.02)	
	Better	11.07 (.78)		24.42 (.60)		42.28 (1.56)	
Self-Care	Same	12.77 (.65)	15.71**	22.84 (.50)	18.76**	44.05 (1.30)	14.20**
	Worse	17.47 (.88)		19.02 (.67)		53.91 (1.75)	
	Yes	12.99 (.49)	6.33*	22.92 (.40)	5.90*	44.54 (1.00)	14.30***

Wellbeing services available	No	16.50 (1.30)		20.19 (1.05)		55.19 (2.64)	
Use wellbeing services	Yes	13.20 (.62)	.34	23.11 (.49)	.48	44.62 (1.25)	.01
	No	12.58 (.87)		22.53 (.68)		44.40 (1.75)	
Concerns about confidentiality	Yes	15.61 (1.05)	5.59*	20.61 (.82)	5.95*	51.11 (2.05)	8.01**
	No	12.84 (.53)		22.84 (.41)		44.60 (1.03)	
Understand the pressure	Yes	12.92 (.52)	5.10*	22.84 (.40)	7.54**	44.77 (1.01)	7.79**
	No	15.78 (1.16)		20.15 (.90)		51.70 (2.27)	

***p < .001, **p < .01, *p<.05