



**ZEVO
HEALTH**

WHITE PAPER

Measuring the Wellbeing of
Content Moderators:
A Clinical Perspective



Dr. Michelle Teo
Clinical Director, Zevo Health

INTRODUCTION

Content moderation is a demanding role that involves the review and management of potentially harmful and distressing material. Ensuring the wellbeing of Content Moderators is crucial, not only for their mental health but also for their productivity and overall job satisfaction. This book explores the current methodologies for measuring the wellbeing of Content Moderators, identifies gaps in these methodologies, and proposes innovative approaches to improve these measurements.

WHY MODERATORS NEED A DIFFERENT WELLBEING MEASUREMENT

Before we look at the types of research we have carried out, we should talk about why Content Moderators need their own measurement of wellbeing:

- 1. Unique Nature of Work** – Content Moderators are routinely exposed to distressing and graphic material, including violence, abuse, and other harmful content. This exposure can lead to significant emotional and psychological distress, which is not typically encountered in other professions so frequently or repetitively.
- 2. Specific Psychological Risks** – The constant exposure to disturbing content and the need to make rapid, high-stakes decisions can lead to high levels of stress and burnout among Content Moderators. Burnout can manifest as emotional exhaustion, depersonalization, and a reduced sense of personal accomplishment. Content Moderators are at risk of developing secondary traumatic stress, which occurs when individuals experience stress symptoms from indirect exposure to trauma through their work. As a chronic experience, this can develop into vicarious trauma which mimics symptoms of post-traumatic stress disorder and also, importantly, can result in profound shifts in worldview.
- 3. Inadequacy of General Wellbeing Measures** – General wellbeing measures do not capture the specific stressors and psychological risks faced by Content Moderators and trauma-based professionals. A dedicated measurement approach is necessary to address the unique challenges of these roles.
- 4. Enhancing Support and Interventions** – Dedicated wellbeing measurement allows for the development of targeted interventions that specifically address the unique stressors faced by Content Moderators and trauma-based professionals. This can include specialized counseling services, peer support programs, and resilience training. Regular and specific wellbeing assessments help in the early identification of mental health issues, allowing for timely intervention. This proactive approach can prevent the escalation of symptoms and reduce the risk of long-term mental health problems.

- 5. Improving Job Performance and Retention** – Employees who feel supported and whose wellbeing is regularly monitored are likely to be more productive. Addressing wellbeing can reduce absenteeism, enhance job performance, and improve overall job satisfaction. Conversely, high levels of stress leading to burnout can result in high turnover rates.
- 6. Legal and Ethical Responsibilities** – Organizations have a legal and ethical duty of care to protect the mental health and wellbeing of their employees. Dedicated measurement and support systems demonstrate a commitment to fulfilling this duty, which can enhance the organization’s reputation and reduce the risk of legal issues.
- 7. Organizational Culture and Morale** – Prioritizing the wellbeing of employees contributes to a positive workplace culture. When employees see and feel that their wellbeing is valued, it fosters a sense of loyalty and commitment to the organization. High morale among Content Moderators and trauma-based professionals can be achieved by ensuring they have the necessary support to manage the emotional and psychological demands of their work.
- 8. Case Studies and Research** – Research and case studies, such as those conducted by Zevo Health, show significant improvements in psychological wellbeing and resilience when dedicated measures are implemented. These studies provide evidence that tailored interventions based on specific wellbeing metrics are effective in supporting Content Moderators and trauma-based professionals. Establishing industry benchmarks for wellbeing in trauma-based roles can help organizations measure their performance against best practices and strive for continuous improvement.

Dedicated measurement of wellbeing for Content Moderators and other trauma-based roles is essential due to the unique and intense nature of their work. It allows for targeted interventions, early identification of mental health issues, rapid intervention, and overall improved job performance and employee retention. By prioritizing the specific wellbeing needs of these professionals, organizations can fulfill their legal and ethical responsibilities, improve workplace culture, and ensure the long-term health and productivity of their workforce.

CHAPTER 1: TRADITIONAL METRICS

Traditionally, the wellbeing of Content Moderators has been assessed using qualitative metrics developed from EAP solutions, including:

- **Utilization and Participation Rates** – Most popular EAP measurement metrics include utilization and participation rates, which track how frequently moderators use support services over different time periods (weekly, monthly, quarterly, annually).
- **Presenting Issues in Counseling** – Another common metric involves documenting the presenting issues during counseling sessions. This helps in identifying prevalent mental health concerns among moderators, such as anxiety, depression, and stress.
- **Breakdown by Demographics** – To add depth to these metrics, data is often broken down by team, shift, location, and onsite status. This helps in understanding how different groups within the organization are affected.
- **Qualitative Insights** – Qualitative data, obtained through anonymization of service engagement disclosures (e.g., from peer group support sessions), focus groups, interviews, and feedback, provides context to the quantitative data. It includes moderators' lived experiences and perceptions of their wellbeing.

Limitations of Traditional Metrics

Despite their usefulness, traditional metrics often lack depth and fail to capture the true meaning and value behind the data. They do not provide sufficient insights into the effectiveness of wellbeing services or the reasons behind the utilization patterns. This makes it challenging to demonstrate the return on investment (ROI) for these services to senior leaders.

CHAPTER 2: PSYCHOLOGICAL HEALTH AND WELLBEING METRICS

In Zevo Health, we have worked to develop a different way of tracking Content Moderator wellbeing, and in this section, we'll talk about the three case studies we have completed.

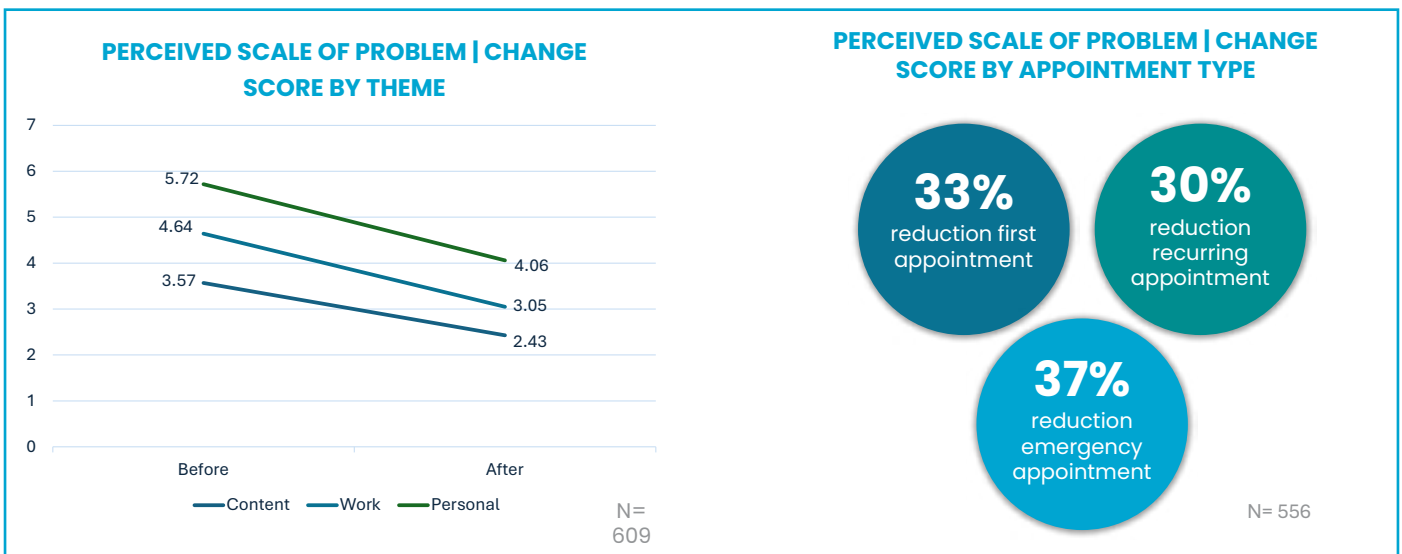
Measure the Effectiveness of Counseling Services

Focused Acceptance and Commitment Therapy (FACT) is a psychological intervention, brief by nature, which utilizes the principles of ACT in a highly condensed format. It has been proven to be effective as a single-session intervention in psychotherapy and counseling settings. As part of the FACT model, clinicians use a scaling question to understand the employee's perceived size of the problem they are facing. This scaling question – asked at the beginning and end of each session – is assessed using a simple Likert scale from 1 to 10 (1 = not a problem, 10 = a very big problem). In Zevo, we call the before and after measurement the change score.

Change scores are calculated to measure the reduction in perceived problem severity within counseling sessions. Although this metric is designed for single-session use, it can also be tracked over a course of sessions. This metric provides a more meaningful assessment of intervention effectiveness compared to participation rates.

Case Study 1: Effectiveness of Counseling Sessions

A study involving 609 counseling sessions showed significant reductions in perceived problem severity:



These results demonstrate the effectiveness of counseling interventions in improving moderators' wellbeing.

Wellbeing Audits

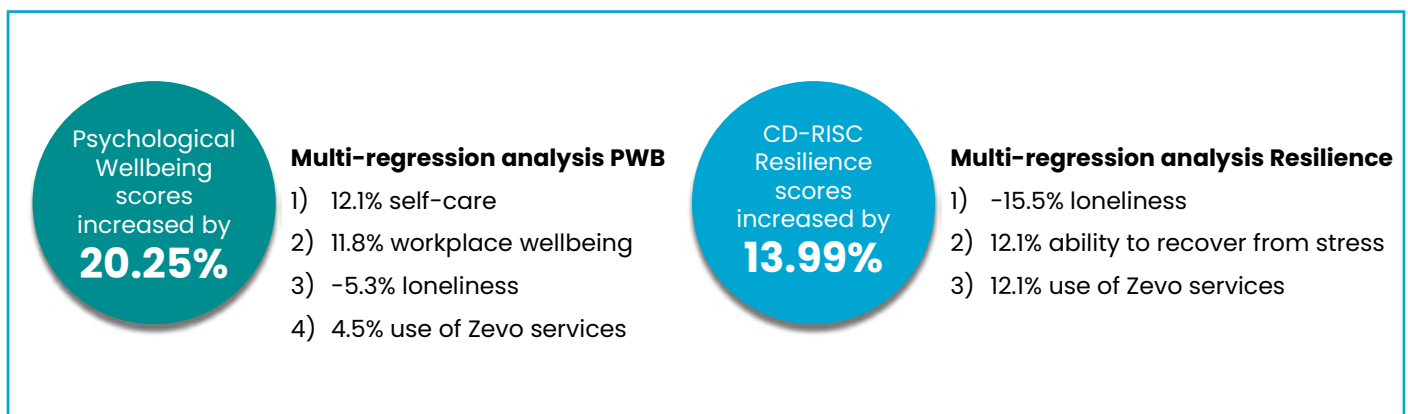
The second case study looks at measuring the psychological health of moderation teams, inclusive of management and moderators. This seeks to address the gaps in measurement around engagement from the emotional involvement and commitment point of view, barriers or reasons for attendance, organizational health, and consistency.

Through repeat measures, using Zevo's audit, we have seen significant improvements in psychological wellbeing scores. However many customers do not engage in repeat assessments therefore we cannot assess change over time. For those that do both an audit and a psychosocial risk assessment (case study 3), we see the greatest benefits across the repeat audit measures.

Wellbeing audits involve repeat measures of psychological wellbeing and resilience. Metrics such as the WHO-5 Psychological Wellbeing Scale and the Connor-Davidson Resilience Scale (CD-RISC) are used to track changes over time.

Case Study 2: Repeat Wellbeing Audit

Two of the metrics we are using are psychological wellbeing and resilience – over time we can see that there have been improvements in the group. We also completed multi-regression analyses during the second audit this year and found:



Multi-regression analyses identified self-care and loneliness as significant predictors of wellbeing and resilience, respectively.

Self-care was the strongest predictor of Psychological Wellbeing (PWB), contributing 12.1% to the prediction of PWB scores. This means that for every

1-point improvement in self-care score, it's predicted there will be a 2.03-point increase in PWB levels. We also found that high loneliness is strongest predictor of low resilience, contributing 15.5% so for every 1-point increase in loneliness scores, it's predicted there will be a 1.5-point decrease in resilience levels.

The importance of this data is understanding the level to which factors can predict improvements or decreases in scores. While it's still imperative to understand the psychological wellbeing score or resilience score of a group of Content Moderators, if we can identify the factors which impact it, then we can target those factors specifically through training and skills development. If loneliness predicts decreased resilience for this team, then we can develop interventions that enhance social connection as a buffer.



Further, we looked at workplace stressors and identified these three as the most impactful across the two audits: performance expectations, organizational communications/structure and reviewing content.

We can derive from this data that unfortunately, performance expectations and reviewing content have become more prevalent stressors than in 2022. Organizational communications, culture or structure has obviously improved, and it has less of an impact as a workplace stressor.

Other Correlations

Wellbeing Domains	Dept A	Dept B	Dept C
PWB & Workplace Wellbeing	Strong	Strong	Medium
PWB & Managerial Support	Medium-Weak	Medium	Weak
Workplace Wellbeing & Managerial Support	Very weak	Very Strong	Strong

Perhaps most interesting was the correlations, finding:

- Very strong correlation between Workplace Wellbeing and Managerial Support in Dept B and strong correlation in Dept C (i.e. staff who report lower stressors also report greater support from management in these departments.)
- Strong correlation between Psychological Wellbeing and Workplace Wellbeing in Dept A & Dept B, medium correlation in Dept C (i.e. staff with high psychological wellbeing tend to be less impacted by stressors).
- Medium strength correlation between Psychological Wellbeing and Managerial Support in Dept B (i.e. those who report good psychological health also feel supported by management.)

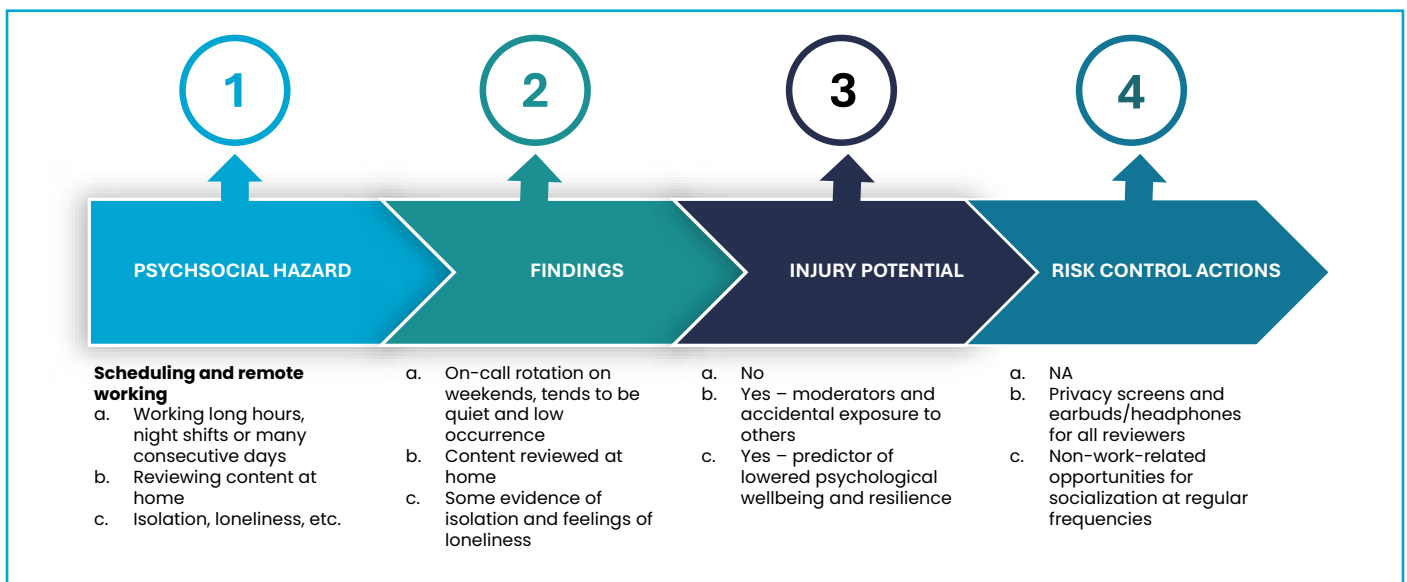
What this case study shows us is that by measuring distinct elements of wellbeing, from general psychological wellbeing to resilience, to workplace stressors and managerial support, we can understand how they interrelate and gives insight into how they may impact one another. This is important as we know that moderators are not only facing egregious content – but we have heard anecdotally, and the research in workplace wellbeing has consistently demonstrated, that factors such as managerial support are key to driving engagement, improved performance, reducing attrition and even enhancing organizational loyalty.

Psychosocial Risk Assessments

Psychosocial risk assessments follow the ISO 45003 framework for psychological health and safety. They involve identifying hazards, assessing risks, implementing control measures, and reassessing.

Case Study 3: Psychosocial Risk Assessment

A study with an audio streaming company identified scheduling, remote working, and content review as significant psychosocial hazards. In focus groups, we uncovered that reviewing content at home and disclosures of feelings of isolation and loneliness had injury potential. Recommendations included privacy screens, earbuds, and socialization opportunities to mitigate risks.



Other areas explored in psychosocial risk assessments include systemic factors such as grievance procedures, sickness absences, job demand and control, organizational change, and exposure to harmful materials, to name a few. Each of these is assessed against their injury potential (or level of risk) and control measures are then put in place and tracked over time.

It is extremely important as an organization to engage in these types of risk assessments as they provide a holistic view of hazards in the workplace – much like occupational health and safety looks at elements like exposed wiring or mold – and addresses them through robust action planning. This not only keeps Content Moderators safe from harm, it ensures that organizations have a clearly defined plan to tackle any major risks and avoid litigation.

CHAPTER 3: A NEW WAY FORWARD

The Zevo Health R&D team has engaged in a pilot research project in which we are co-designing a measure that looks at subjective wellbeing, work-life satisfaction, and perceived organizational support for wellbeing. We chose these factors based upon years of data collection with our existing customers, and the lived experience of clinicians working directly with moderation teams. These factors were identified to be the most impactful on moderators.

The research is only three short questions, intended to be cross-culturally relevant and minimize survey fatigue, while deriving meaningful information. Our first pilot went through four stages last year:

- **Stage 1:** development of measure including cross-cultural considerations, descriptors for emotions and cognitions, color palettes and imagery, and language tailoring
- **Stage 2:** eight-week pilot with internal Zevo employees
- **Stage 3:** facilitation of focus groups to identify how individuals interpreted the questions, barriers or reasons for (un)responsiveness, cross-cultural differences, etc.
- **Stage 4:** refinement phase – we added one more question!

Pilot 2 is running now with four different groups: a government-funded organization, an NGO, dating app moderation team, and our own internal staff who are in T&S wellbeing

Our hypotheses are simple:

Subjective Wellbeing – Improved Subjective Wellbeing over time with regularly scheduled wellbeing supports in place; correlated with perceived org support for wellbeing. The project involved regular surveys and focus groups to refine the measurement approach.

- **Work-life satisfaction** – Correlated with perceived organizational support for wellbeing and SWB
- **Perceived org support** – More regular access to wellbeing supports improve perceived organizational support; correlate with work-life satisfaction and SWB.

We are excited to see the results of pilot 2 of this research study.

FUTURE DIRECTIONS

As researchers and wellbeing service providers, we aspire to a handful of innovations in measuring moderator wellbeing. We would hope to see the T&S industry globally move away from individual measurement towards organizational health assessments, supported by robust and compliant data sharing between departments that can then lead to industry benchmarking.

We believe the industry needs to move away from absolute figures of utilization towards measuring the effectiveness of services on employee wellbeing and engagement, which will allow for more nuanced understanding of the challenges faced by moderation teams and appropriateness of psychological interventions. We are also eager to see organizations explore the development of specific psychometric scales relevant to the industry and Content Moderators' roles.

And finally, we continue to hope that everyone aims to reduce survey fatigue when collecting metrics from moderation teams, moving towards more short, sweet, and meaningful measurement options.

Content Moderator Wellbeing Index

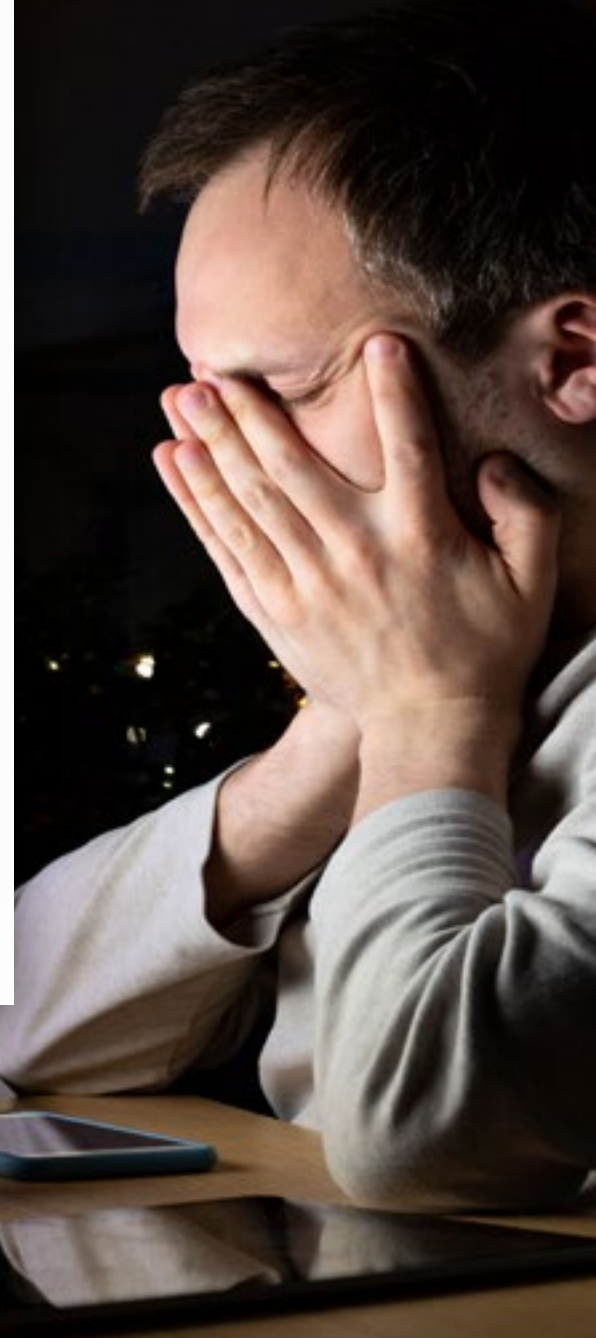
Next, we plan to focus on developing a specific wellbeing index for Content Moderators, similar to the Well-Being Index[®] for nurses and physicians. This will help us provide more targeted insights.

CONCLUSION

Measuring the wellbeing of Content Moderators is a complex task that requires a combination of quantitative and qualitative approaches. Traditional metrics provide a baseline, but more innovative and holistic methods are needed to capture the true state of wellbeing and effectiveness of interventions. By adopting a comprehensive measurement framework, organizations can better support their content moderation teams, ensuring their mental health and enhancing their productivity and job satisfaction.

References

- Zevo Health's case studies and research on moderator wellbeing
- ISO 45003 framework for psychological health and safety
- WHO-5 Psychological Wellbeing Scale
- Connor-Davidson Resilience Scale (CD-RISC)





ZEVO HEALTH

Empowering the psychological health and
wellbeing of content moderation teams



www.zevohealth.com



hello@zevohealth.com