

When Care Becomes Harm: Sexual Exploitation, Abuse and Harassment in Mental Health Care: Advancing a Mental Health Safeguarding Framework from Kenya to Global Practice

Emily Gideon

Psycho Africa Solutions

Corresponding author email: gideon.emilysam@gmail.com;
emilygideon@psychoafrica.co.ke

Accepted: 21 July 2026 || Published: 18 September 2026

Abstract

Background: Sexual exploitation, abuse, and harassment (SEA/SH) constitute serious violations of professional ethics, patient safety, and human rights. Although safeguarding has become integral to healthcare, humanitarian, and education sectors, its application within mental health care remains underdeveloped despite the unique relational and ethical dynamics of therapeutic practice.

Objective: This review advances Mental Health Safeguarding as an emerging interdisciplinary field and proposes the Mental Health SEA/SH Safeguarding Framework (MHSSF) to strengthen prevention, early identification, response, accountability, and organisational learning. Using Kenya as a contextual lens, the paper demonstrates how safeguarding principles can be operationalized within mental health systems while generating globally applicable insights.

Methods: A critical narrative review was conducted using multidisciplinary literature from mental health, safeguarding, professional ethics, gender-based violence, human rights, organisational governance, and implementation science. Evidence was synthesised through thematic analysis and conceptual integration to develop an evidence-informed safeguarding framework.

Results: Existing literature largely addresses professional boundaries, workplace harassment, ethics, and patient safety separately, leaving safeguarding within mental health conceptually fragmented. The review introduces four interrelated concepts, Dynamic Vulnerability, Therapeutic Power Asymmetry, Boundary Erosion, and Organisational Safeguarding Capacity, which are integrated into the MHSSF. The framework provides a systems-based approach for strengthening safeguarding across governance, policy, professional education, clinical practice, regulation, and organisational leadership.

Conclusion: Safeguarding should be recognised as a core dimension of quality mental health care rather than merely a response to professional misconduct. The MHSSF provides a conceptual foundation for strengthening trauma-informed, rights-based, and accountable mental health systems while supporting future research, policy development, and implementation across diverse settings.

Keywords: *Mental health safeguarding, Sexual exploitation, Sexual abuse, Sexual harassment, Professional ethics, Therapeutic relationships, Human rights, Kenya.*

How to Cite: Gideon, E. (2026). When Care Becomes Harm: Sexual Exploitation, Abuse and Harassment in Mental Health Care: Advancing a Mental Health Safeguarding Framework from Kenya to Global Practice. *Journal of Sociology, Psychology and Religious Studies*, 6(6), 57-79.

1. Introduction

Mental health care occupies a unique position within healthcare because therapeutic effectiveness depends on relationships characterised by trust, emotional engagement, and sustained interpersonal interaction. Mental health professionals routinely access deeply personal information concerning trauma, abuse, relationships, suicidal ideation, and psychological distress, making the therapeutic alliance both the principal mechanism of healing and a potential source of vulnerability when professional boundaries are violated (Bordin, 1979; Norcross & Lambert, 2018).

The therapeutic relationship consistently predicts positive treatment outcomes regardless of therapeutic orientation (Wampold & Imel, 2015). However, these same relationships create inherent asymmetries of knowledge, authority, and emotional influence that require strong ethical safeguards. Consequently, safeguarding in mental health extends beyond technical competence to encompass professional boundaries, accountability, organisational governance, and protection of those receiving care (American Psychological Association [APA], 2017).

Sexual exploitation, abuse, and harassment (SEA/SH) are among the most serious violations of professional ethics and patient safety. Such misconduct undermines dignity, retraumatises survivors, erodes confidence in health systems, and compromises therapeutic recovery. Global safeguarding initiatives increasingly recognise that abuse arises where unequal power intersects with inadequate institutional protections, leading to the adoption of safeguarding frameworks across humanitarian, education, disability, and healthcare sectors (Inter-Agency Standing Committee [IASC], 2015; World Bank, 2020). Despite these advances, safeguarding within mental health services remains comparatively underdeveloped.

Current scholarship largely examines professional boundaries, ethics, workplace harassment, trauma-informed care, patient rights, and institutional misconduct as separate domains. Few studies integrate these perspectives to explain how therapeutic relationships, emotional dependency, organisational culture, and governance interact to create safeguarding risks. This gap is particularly significant in low- and middle-income countries, where mental health systems are expanding amid workforce shortages, limited resources, and evolving regulatory frameworks.

This review argues that safeguarding within mental health requires a systems perspective. Rather than viewing SEA/SH solely as individual misconduct, safeguarding failures should be understood as arising from interactions between changing vulnerability, therapeutic power, professional boundaries, organisational culture, and governance structures. Prevention therefore requires coordinated interventions operating at individual, relational, organisational, and policy levels.

To address this gap, the paper proposes the Mental Health SEA/SH Safeguarding Framework (MHSSF). The framework integrates safeguarding science, professional ethics, trauma-informed care, organisational governance, and human rights into a unified model comprising four interrelated constructs: Dynamic Vulnerability, Therapeutic Power Asymmetry, Boundary Erosion, and Organisational Safeguarding Capacity. Together, these concepts provide an evidence-informed framework for strengthening prevention, early identification, response, accountability, and organisational learning within mental health systems.

Using Kenya as a contextual example while drawing on international evidence, this review positions Mental Health Safeguarding as an emerging interdisciplinary field capable of informing policy, research, professional education, and service delivery globally.

2. Methodology

This study employed a critical narrative review to examine the intersection of sexual exploitation, abuse and harassment (SEA/SH), therapeutic relationships, professional ethics, safeguarding systems, and mental health care. A narrative review was appropriate because the topic spans multiple disciplines and seeks conceptual integration rather than quantitative synthesis (Grant & Booth, 2009; Greenhalgh et al., 2018).

Relevant literature was identified through searches of PubMed, PsycINFO, Scopus, Web of Science, and Google Scholar, supplemented by policy documents and professional guidance from organisations including the World Health Organization (WHO), Inter-Agency Standing Committee (IASC), World Bank, and international professional associations. Literature addressed safeguarding, professional boundaries, trauma-informed care, organisational governance, patient safety, workplace harassment, human rights, and mental health systems.

Sources were selected for relevance to safeguarding in mental health settings and included peer-reviewed research, theoretical publications, professional ethical codes, policy frameworks, and international guidance. Thematic synthesis focused on four analytical questions:

1. How is vulnerability conceptualised within mental health care?
2. How do therapeutic relationships create both opportunities for healing and safeguarding risks?
3. How do organisational culture and governance influence safeguarding outcomes?
4. What conceptual framework can integrate these dimensions into a coherent safeguarding model?

Conceptual synthesis identified four recurring constructs: Dynamic Vulnerability, Therapeutic Power Asymmetry, Boundary Erosion, and Organizational Safeguarding Capacity. These constructs were integrated into the proposed Mental Health SEA/SH Safeguarding Framework (MHSSF), which is presented as a theoretically informed model requiring future empirical validation.

The review adopts a survivor-centered, rights-based, and trauma-informed perspective while recognizing that safeguarding is a shared responsibility of practitioners, organizations, regulators, and health systems. Although the review does not estimate the prevalence of SEA/SH in mental health settings, it synthesizes evidence from adjacent fields to develop a comprehensive conceptual framework applicable across diverse mental health contexts.

3. Sexual Exploitation, Abuse and Harassment in Mental Health Care: An Under-recognised Safeguarding Challenge

3.1 Understanding SEA/SH in Mental Health Care

Sexual exploitation, abuse, and harassment (SEA/SH) are serious violations of human rights, professional ethics, and patient safety. Although safeguarding has gained prominence across healthcare, humanitarian, education, and social service sectors, its application within mental health care remains comparatively underdeveloped (Inter-Agency Standing Committee [IASC], 2015; World Bank, 2020). Mental health services present distinctive safeguarding challenges because therapeutic relationships are built upon trust, confidentiality, emotional disclosure, and prolonged professional engagement.

Unlike many healthcare encounters, mental health practitioners gain intimate knowledge of service users' personal histories, trauma, relationships, fears, and emotional experiences. These therapeutic interactions are essential for recovery but simultaneously create unequal distributions of knowledge, authority, and influence. Consequently, safeguarding within mental health cannot be limited to responding to misconduct after it occurs; it must proactively address the conditions that enable exploitation and abuse.

SEA/SH within mental health extends beyond inappropriate sexual relationships between clinicians and patients. Safeguarding concerns may arise across practitioner–patient, supervisor–trainee, educator–student, employer–employee, and colleague relationships where professional authority and dependency coexist. Effective safeguarding therefore requires recognising the multiple forms of power operating throughout mental health systems.

3.2 SEA/SH Within the Context of Gender-Based Violence

SEA/SH should be understood as a specific manifestation of gender-based violence (GBV), characterised by the abuse of authority, trust, or dependency for sexual purposes. While GBV encompasses a broad spectrum of harmful behaviours rooted in gender inequality, SEA/SH specifically occurs where individuals misuse positions of power to exploit those who depend on them for care, employment, education, or protection (World Health Organization [WHO], 2013).

Mental health settings present unique vulnerabilities because many individuals seeking care have previously experienced sexual violence, childhood abuse, intimate partner violence, or other traumatic events (Herman, 1992). Therapeutic relationships often become central sources of safety, validation, and recovery. However, violations of professional boundaries within these relationships may retraumatise service users, worsen psychological distress, reduce treatment engagement, and undermine trust in mental health services.

Importantly, exploitation in mental health settings does not always involve physical force. Manipulation, grooming, coercion, emotional dependency, inappropriate self-disclosure, or misuse of professional authority may progressively erode boundaries before overt abuse occurs. Recognising these subtle processes is essential for effective prevention.

3.3 Manifestations of SEA/SH in Mental Health Systems

SEA/SH may occur through a continuum of behaviours ranging from inappropriate communication to severe criminal misconduct. Common manifestations include:

- Sexual relationships between practitioners and current service users.

- Sexualised comments, jokes, or inappropriate communication.
- Grooming through excessive personal disclosure or preferential treatment.
- Exploitation of supervisory or educational authority.
- Sexual harassment among staff or trainees.
- Emotional manipulation leading to dependency.
- Retaliation against individuals reporting misconduct.

Professional ethical codes consistently prohibit sexual relationships with current clients because genuine consent is compromised by the inherent imbalance of authority within therapeutic relationships (American Psychological Association [APA], 2017; Gutheil & Gabbard, 1993). Nevertheless, safeguarding should not focus solely on completed acts of misconduct. Research demonstrates that serious ethical violations are frequently preceded by progressive boundary violations, secrecy, and increasing personal involvement (Gabbard, 2016). Early recognition of these warning signs enables organisations to intervene before significant harm occurs.

3.4 Why Mental Health Requires a Distinct Safeguarding Approach

Several characteristics distinguish mental health care from other healthcare environments and increase safeguarding complexity.

First, **therapeutic relationships** require sustained emotional engagement. Service users often disclose highly sensitive experiences over extended periods, creating strong interpersonal bonds that can increase dependence on practitioners.

Second, **professional authority** extends beyond diagnosis and treatment. Mental health professionals influence access to services, legal assessments, medication, rehabilitation programmes, and psychosocial support. This authority creates opportunities for both therapeutic benefit and potential misuse.

Third, **confidentiality** limits external observation of clinical encounters. While confidentiality is essential for effective therapy, it may also reduce opportunities to detect inappropriate behaviour if organisations lack robust supervision and reporting systems.

Finally, mental health services often involve individuals experiencing periods of psychological distress, trauma, or impaired coping capacity. Safeguarding therefore requires particular attention to changing vulnerability and ethical management of therapeutic relationships rather than relying solely on traditional professional conduct procedures.

3.5 Limitations of Existing Safeguarding Approaches

Current safeguarding approaches have several limitations in mental health settings. Many frameworks remain reactive, emphasising investigation after allegations rather than prevention through early identification of relational risks. Professional ethical codes primarily define individual responsibilities but provide limited guidance regarding organisational accountability, leadership, supervision, and institutional culture. Furthermore, many safeguarding models were developed within general healthcare or humanitarian contexts and do not adequately address the distinctive psychological processes characterising mental health care, including emotional dependency, transference, countertransference, and prolonged therapeutic engagement.

Evidence from low- and middle-income countries is particularly limited. Expanding mental health services often occurs alongside workforce shortages, limited supervision, evolving regulations, and inadequate reporting systems, underscoring the importance of integrated safeguarding strategies.

These limitations demonstrate the need for a mental health-specific safeguarding framework that addresses interactions among vulnerability, therapeutic relationships, organisational governance, and professional ethics.

4. Dynamic Vulnerability: Reframing Vulnerability Within Mental Health Safeguarding

4.1 Moving Beyond Static Understandings of Vulnerability

Vulnerability is central to safeguarding, healthcare ethics, and human rights. Traditionally, vulnerability has been associated with characteristics such as illness, disability, age, or reduced decision-making capacity (Levine et al., 2004). Within mental health services, people experiencing psychological distress are often categorised as inherently vulnerable because of their clinical conditions or dependence on professional care.

Although recognising vulnerability is essential for protection, static classifications can inadvertently reinforce paternalistic assumptions and overlook individuals' resilience, autonomy, and recovery. Contemporary scholarship increasingly conceptualizes vulnerability as relational and context-dependent, arising from interactions among personal circumstances, social environments, institutions, and relationships rather than as an intrinsic personal characteristic (Fineman, 2008; Luna, 2009).

This perspective is particularly relevant to mental health safeguarding because safeguarding risks fluctuate throughout the therapeutic journey.

4.2 Dynamic Vulnerability as a Safeguarding Concept

This review introduces Dynamic Vulnerability as the first pillar of the Mental Health SEA/SH Safeguarding Framework (MHSSF).

Dynamic Vulnerability refers to the changing susceptibility of individuals to safeguarding risks resulting from interactions among clinical status, trauma history, therapeutic engagement, social circumstances, and organisational conditions.

Unlike traditional approaches, Dynamic Vulnerability recognises that safeguarding needs change over time. Psychological crises, recovery, social isolation, trauma exposure, disability, communication barriers, and organisational environments may each increase or reduce safeguarding risk. Importantly, recognising vulnerability does not imply diminished autonomy. Individuals may simultaneously experience vulnerability while retaining significant decision-making capacity and personal resilience.

This approach shifts safeguarding from categorising "vulnerable people" to identifying situations that create vulnerability.

4.3 Mental Health Conditions and Changing Safeguarding Needs

Mental health conditions influence safeguarding in diverse ways. During periods of acute depression, psychosis, substance misuse, severe anxiety, or suicidal crisis, individuals may experience temporary difficulties in judgement, emotional regulation, or help-seeking. However, diagnosis alone does not determine vulnerability. Many individuals living with mental illness function independently and actively participate in treatment decisions.

Safeguarding therefore requires continuous assessment rather than assumptions based on diagnosis. A person's vulnerability may increase during crisis yet decrease substantially during recovery. Organisations should consequently develop flexible safeguarding responses that adapt to changing circumstances while respecting autonomy and recovery-oriented care.

4.4 Trauma, Attachment and Relational Vulnerability

Many individuals accessing mental health services have experienced adverse childhood experiences, interpersonal violence, conflict, displacement, or sexual abuse (Herman, 1992). These experiences may influence trust, attachment, expectations of care, and responses to authority figures.

Therapeutic relationships often provide opportunities for healing through empathy, validation, and emotional safety. However, the same relational significance may increase susceptibility to exploitation if practitioners misuse professional authority or fail to maintain clear boundaries.

Trauma-informed safeguarding therefore requires organisations to recognise that previous victimisation can interact with therapeutic relationships without assuming that survivors are inherently incapable of protecting themselves.

4.5 Organisational Responsibility for Dynamic Vulnerability

Dynamic Vulnerability shifts the responsibility for safeguarding from individuals to systems.

Rather than asking “Who is vulnerable?”, organisations should ask:

- Under what circumstances does vulnerability increase?
- Which organisational factors amplify or reduce safeguarding risks?
- How can services adapt to changing safeguarding needs?

This systems perspective requires organisations to evaluate supervision, accessibility, reporting mechanisms, professional boundaries, workforce competence, and organisational culture alongside individual clinical assessments.

Safeguarding thus becomes a shared responsibility involving practitioners, organisations, professional regulators, and policymakers.

4.6 Implications for Mental Health Practice

Dynamic Vulnerability has several practical implications.

Clinicians should integrate safeguarding assessment into routine care rather than treating it as a one-time procedure. Professional education should prepare practitioners to recognise how trauma, therapeutic relationships, and contextual factors influence safeguarding risk. Organisations should strengthen supervision, reflective practice, reporting systems, and continuous monitoring to respond to changing vulnerability. Researchers should develop measures that can assess vulnerability dynamically across different mental health settings.

4.7 Conceptual Contribution

Dynamic Vulnerability provides the conceptual foundation of the MHSSF by reframing vulnerability as fluid, relational, and context-dependent. Rather than labelling individuals as permanently vulnerable, the concept recognises that safeguarding risks emerge through interactions between clinical experiences, therapeutic relationships, and organisational systems. This perspective strengthens ethical, person-centred, and recovery-oriented mental health practice while establishing the basis for understanding the second major construct of the framework—Therapeutic Power Asymmetry.

5. Therapeutic Power Asymmetry: The Paradox of Healing and Harm in Mental Health Relationships

5.1 Understanding Power in Therapeutic Relationships

Power is an inherent feature of all professional relationships, but it assumes particular significance in mental health care because treatment depends on trust, emotional disclosure, and sustained interpersonal engagement. Mental health professionals possess specialised knowledge, clinical expertise, institutional authority, and responsibility for decisions that influence diagnosis, treatment, access to services, and recovery. These forms of authority are indispensable for effective care; however, they also create unequal relationships that require strong ethical safeguards (Bordin, 1979; Norcross & Lambert, 2018).

Unlike many healthcare interactions, therapeutic relationships involve discussions of trauma, identity, relationships, sexuality, grief, and psychological distress. Service users often disclose experiences that they have never shared with others, making therapists trusted custodians of highly sensitive information. This relational depth enhances treatment effectiveness but simultaneously increases the potential for exploitation if professional authority is misused.

Safeguarding should therefore recognise that therapeutic power is neither inherently positive nor inherently harmful. Rather, its ethical value depends on how it is exercised, regulated, and supported through professional accountability and organisational governance.

5.2 Therapeutic Power Asymmetry

This review introduces **Therapeutic Power Asymmetry** as the second foundational construct of the Mental Health SEA/SH Safeguarding Framework (MHSSF).

Therapeutic Power Asymmetry refers to the inherent imbalance of authority, knowledge, emotional influence, and institutional legitimacy between mental health practitioners and those receiving care.

This asymmetry operates through four interconnected dimensions.

Knowledge asymmetry arises because clinicians possess specialised expertise regarding diagnosis, treatment, and psychological functioning. Service users frequently rely on this expertise to understand their conditions and make treatment decisions.

Institutional asymmetry reflects the authority conferred through professional registration, licensure, organisational position, and legal responsibilities. Practitioners act not only as individuals but also as representatives of healthcare institutions and professional bodies.

Emotional asymmetry develops because therapeutic relationships require service users to disclose highly personal experiences while practitioners maintain professional objectivity. This unequal exchange of personal information creates significant relational influence.

Dependency asymmetry occurs when individuals experiencing psychological distress depend on practitioners for emotional support, crisis intervention, referrals, certification, or continued access to services. During periods of crisis, this dependency may temporarily increase, reinforcing the importance of professional integrity.

Together, these dimensions explain why mental health professionals bear heightened ethical responsibilities beyond those found in ordinary interpersonal relationships.

5.3 Therapeutic Power: A Resource for Healing

Professional authority should not be viewed solely as a safeguarding risk. Therapeutic power is essential for effective mental health care because it enables clinicians to assess risk, provide psychological interventions, advocate for service users, coordinate multidisciplinary care, and support recovery. Without legitimate professional authority, practitioners would be unable to fulfil these responsibilities effectively.

The safeguarding challenge therefore lies not in reducing therapeutic authority but in ensuring that it is exercised transparently, ethically, and within clearly defined professional boundaries. Organisations should foster environments in which authority is consistently accompanied by accountability, reflective practice, and clinical supervision.

5.4 Relational Complexity: Transference and Countertransference

Mental health practice involves relational processes that distinguish it from many other healthcare settings. Two concepts are particularly relevant: **transference** and **countertransference** (Gelso & Hayes, 2007).

Transference refers to the unconscious transfer of feelings, expectations, or relational patterns from previous significant relationships onto the therapist. Countertransference describes practitioners' emotional responses to service users and the influence these reactions may have on therapeutic decision-making.

When recognised through reflective practice and supervision, these processes enhance therapeutic understanding. However, when ignored or poorly managed, they may contribute to blurred boundaries, emotional over-involvement, or inappropriate dependency. Consequently, supervision should address not only clinical interventions but also the relational dynamics that influence safeguarding.

5.5 Boundary Erosion as the Pathway to Safeguarding Failure

Therapeutic Power Asymmetry explains why safeguarding risks exist; however, it does not fully explain how serious misconduct develops. This review proposes Boundary Erosion as the mechanism linking legitimate professional authority to safeguarding failure.

Boundary erosion refers to the gradual weakening of professional boundaries through repeated departures from accepted ethical standards. These departures may initially appear insignificant or even benevolent but progressively normalise inappropriate behaviour.

Common indicators include:

- Excessive personal self-disclosure by practitioners;
- Communication outside established professional channels;
- Favouritism or "special" therapeutic relationships;
- Blurred social and professional roles;
- Secrecy surrounding interactions;
- Emotional dependency encouraged by practitioners; and
- Rationalisation of inappropriate conduct.

Research consistently demonstrates that major ethical violations are rarely isolated incidents; rather, they are often preceded by incremental boundary shifts that go unchallenged (Gabbard, 2016). Recognising these early warning signs enables organisations to intervene before exploitation or abuse occurs.

5.6 Safeguarding Implications

Therapeutic Power Asymmetry has several practical implications.

First, practitioners must acknowledge that consent in therapeutic relationships cannot be interpreted the same way as consent between equals, because professional authority inherently shapes decision-making.

Second, clinical supervision should function as both a professional development and safeguarding mechanism, enabling practitioners to reflect on emotional responses, ethical uncertainty, and emerging boundary concerns.

Third, organisations should create cultures where ethical dilemmas can be openly discussed without fear of blame, thereby strengthening prevention rather than relying solely on disciplinary action after harm occurs.

Finally, professional education should explicitly address power, privilege, professional boundaries, and relational ethics as core safeguarding competencies.

5.7 Conceptual Contribution

Therapeutic Power Asymmetry represents the second pillar of the MHSSF. It demonstrates that the same authority essential for therapeutic healing also creates obligations for protection and accountability. By shifting attention from individual misconduct alone to the responsible governance of professional authority, the concept provides a stronger theoretical foundation for safeguarding within mental health systems.

6. Organisational Safeguarding Capacity: Building Systems That Prevent Harm

6.1 Safeguarding Beyond Individual Accountability

Although individuals are responsible for unethical or abusive behaviour, safeguarding failures rarely occur in isolation. Evidence from patient safety and organisational governance demonstrates that adverse events often reflect broader weaknesses in leadership, supervision, organisational culture, communication, and accountability (Reason, 2000; Vincent, 2010).

Mental health services present particular organisational challenges because clinical encounters are frequently confidential and occur in private settings. Consequently, safeguarding depends

not only on individual professionalism but also on organisational systems capable of preventing, detecting, and responding to emerging risks.

The key question therefore extends beyond “Who caused harm?” to include “What organisational conditions allowed harm to occur or remain undetected?”

6.2 Organisational Safeguarding Capacity

The third pillar of the MHSSF is Organisational Safeguarding Capacity, defined as:

The ability of a mental health organisation to prevent, identify, respond to, and learn from safeguarding risks through effective leadership, governance, workforce competence, reporting systems, supervision, and continuous organisational improvement.

Possessing safeguarding policies alone is insufficient. Effective safeguarding depends upon implementation, leadership commitment, and organisational culture.

Organisations with limited safeguarding capacity often exhibit:

- Inadequate supervision;
- Unclear reporting pathways;
- Insufficient workforce training;
- Fear of retaliation among staff;
- Weak leadership commitment; and
- Cultures that discourage disclosure.

By contrast, organisations with strong safeguarding capacity integrate safeguarding into everyday clinical governance and quality improvement.

6.3 Core Components of Organisational Safeguarding Capacity

Leadership and Governance

Leadership establishes organisational priorities, allocates resources, and shapes institutional culture. Safeguarding should therefore be embedded within governance structures rather than delegated exclusively to ethics committees or human resource departments.

Senior leaders should ensure clear safeguarding responsibilities, regular monitoring, transparent accountability, and continuous review of safeguarding performance.

Policies and Reporting Systems

Safeguarding policies should be translated into practical procedures that are accessible to staff, service users, and families. Effective systems include confidential reporting pathways, survivor-centred responses, transparent investigations, documentation standards, and referral mechanisms.

Mental health organisations should adapt safeguarding procedures to reflect the unique relational characteristics of therapeutic practice.

Workforce Competence

Safeguarding competence extends beyond legal compliance. Mental health professionals require ongoing education regarding professional boundaries, ethical decision-making, trauma-

informed care, management of therapeutic power, and appropriate responses to disclosures of abuse.

Continuous professional development and reflective learning should reinforce safeguarding throughout practitioners' careers.

Clinical Supervision

Clinical supervision is among the most effective organisational safeguards available within mental health care. Beyond improving clinical competence, supervision provides structured opportunities to examine ethical dilemmas, relational complexity, emotional responses, and emerging boundary concerns before they escalate into safeguarding failures.

Reporting Culture and Psychological Safety

Effective safeguarding depends on organisational cultures that encourage individuals to raise concerns without fear of punishment or retaliation. Psychological safety—the shared belief that concerns can be expressed openly—is essential for organisational learning and prevention (Edmondson, 1999).

Mental health organisations should actively address barriers such as professional hierarchy, institutional defensiveness, conflicts of interest, and fear of damaging careers.

Survivor-Centred Response

Responses to SEA/SH should prioritise safety, dignity, confidentiality, and informed choice. Organisations should ensure respectful handling of disclosures, protection from retaliation, access to psychosocial and legal support, and transparent investigative processes.

Safeguarding systems that prioritise institutional reputation over survivor wellbeing risk perpetuating harm and undermining public trust.

6.4 Organisational Culture as a Determinant of Safety

Policies operate within organisational cultures that influence whether safeguarding concerns are recognised or ignored. Positive safeguarding cultures promote openness, ethical leadership, accountability, learning, and respect for service users' rights.

Conversely, cultures characterised by silence, protection of influential individuals, dismissal of complaints, or normalisation of boundary violations significantly increase safeguarding risk. Sustainable safeguarding therefore requires cultural transformation alongside procedural reform.

6.5 Integrating Organisational Safeguarding Capacity into the MHSSF

Within the MHSSF, Organisational Safeguarding Capacity links individual and relational safeguarding risks to institutional responsibility.

The framework proposes that:

- **Dynamic Vulnerability** identifies changing safeguarding needs.
- **Therapeutic Power Asymmetry** explains inherent relational inequalities.
- **Boundary Erosion** describes the progression of safeguarding risk.
- **Organisational Safeguarding Capacity** determines whether organisations prevent, identify, or permit harm.

This integrated perspective positions organisations as active participants in safeguarding rather than passive settings in which misconduct occurs.

6.6 Implications for Policy and Practice

The concept has important implications across mental health systems.

Mental health organisations should integrate safeguarding into governance, quality assurance, accreditation, and risk management processes. Professional regulators should recognise safeguarding competence as a core requirement for licensure and continuing professional development. Universities and training institutions should embed safeguarding throughout mental health curricula rather than treating it as an isolated ethical topic. Policymakers should incorporate safeguarding indicators into national mental health strategies and health system performance frameworks.

6.7 Conceptual Contribution

Organisational Safeguarding Capacity completes the third major pillar of the MHSSF by demonstrating that safeguarding is fundamentally a systems responsibility. Safe therapeutic environments are created not only through ethical practitioners but through organisations that cultivate accountability, transparency, reflective practice, and continuous learning. Integrating organisational governance with professional ethics strengthens prevention and positions safeguarding as a defining characteristic of high-quality mental health care rather than a reactive response to misconduct.

7. The Mental Health SEA/SH Safeguarding Framework (MHSSF): An Integrated Systems Model

7.1 From Conceptual Analysis to an Integrated Framework

The preceding sections demonstrate that safeguarding within mental health cannot be explained through a single perspective. Existing approaches have largely focused on professional misconduct, ethical codes, workplace harassment, or legal accountability independently. While these perspectives are valuable, they do not sufficiently explain how individual vulnerability, therapeutic relationships, organisational culture, and governance interact to produce safeguarding outcomes.

This review addresses this gap through the Mental Health SEA/SH Safeguarding Framework (MHSSF), a systems-based conceptual model that integrates safeguarding science, professional ethics, trauma-informed care, human rights, and organisational governance into a unified approach for preventing and responding to sexual exploitation, abuse, and harassment within mental health settings.

Rather than treating safeguarding as an isolated organisational function, the MHSSF conceptualises it as a continuous process operating simultaneously at individual, relational, organisational, and system levels. The framework recognises that safeguarding failures rarely result from a single event; instead, they emerge through interactions between changing vulnerability, therapeutic power, professional boundaries, and organisational capacity.

Accordingly, the MHSSF is organised around four interconnected constructs:

1. Dynamic Vulnerability
2. Therapeutic Power Asymmetry
3. Boundary Erosion
4. Organisational Safeguarding Capacity

Together, these constructs explain both how safeguarding risks develop and how they can be prevented through coordinated organisational and professional action.

7.2 The Structure of the MHSSF

The MHSSF adopts an ecological systems perspective in which safeguarding is understood as the product of continuous interaction between individuals, therapeutic relationships, organisations, and the wider health system.

Individual Level: Dynamic Vulnerability

Safeguarding begins with recognising that vulnerability is dynamic rather than fixed. Clinical status, trauma history, social support, recovery, and contextual circumstances influence an individual's susceptibility to exploitation at different stages of care. Consequently, safeguarding assessments should be continuous and responsive, not based solely on diagnostic labels or assumptions about capacity.

Relational Level: Therapeutic Power Asymmetry

The therapeutic relationship represents the central mechanism through which both healing and harm may occur. Professional authority, emotional influence, and institutional legitimacy are necessary for effective care but require careful ethical management. The framework therefore positions Therapeutic Power Asymmetry as an unavoidable feature of mental health practice that demands accountability, not avoidance.

Boundary Erosion

Boundary Erosion functions as the connecting process between legitimate professional authority and safeguarding failure. Rather than viewing misconduct as isolated incidents, the framework suggests that exploitation often develops through a gradual weakening of professional boundaries. Early recognition of inappropriate relational changes therefore becomes a central safeguarding strategy.

Organisational Level: Organisational Safeguarding Capacity

Ultimately, safeguarding outcomes depend upon the organisation within which therapeutic relationships occur. Leadership commitment, governance, supervision, workforce competence, reporting systems, organisational culture, and continuous learning determine whether safeguarding risks are prevented, detected early, or allowed to escalate.

These four constructs interact continuously throughout the delivery of mental health care, emphasising that safeguarding is not an isolated intervention but an ongoing organisational responsibility.

7.3 Guiding Principles of the MHSSF

The MHSSF is underpinned by six mutually reinforcing principles that support safe, ethical, and accountable mental health care.

Prevention Before Harm

Safeguarding should prioritise prevention rather than relying primarily on disciplinary responses after abuse has occurred. Early recognition of boundary concerns, effective supervision, workforce education, and organisational vigilance are essential components of prevention.

Survivor-Centred and Rights-Based Practice

Safeguarding interventions should uphold dignity, autonomy, confidentiality, informed participation, and non-discrimination. Individuals reporting SEA/SH should have access to appropriate psychological, legal, medical, and organisational support consistent with international human rights principles.

Therapeutic Integrity

Safeguarding should strengthen, not weaken, therapeutic relationships. Ethical use of professional authority, trust, empathy, and clinical competence remains fundamental to effective mental health care and should be protected through robust safeguarding systems.

Organisational Accountability

Safeguarding is a collective organisational responsibility. Leaders, managers, clinicians, educators, and regulators share responsibility for creating environments that prevent exploitation, encourage reporting, and promote accountability.

Trauma-Informed Systems

Mental health organisations should recognise the widespread impact of trauma among service users, families, and staff. Policies and responses should minimise re-traumatisation while promoting recovery, empowerment, and psychological safety.

Continuous Learning

Safeguarding systems should be regularly evaluated through monitoring, supervision, quality improvement, incident review, and workforce development. Organisational learning enables institutions to adapt to emerging risks and strengthen safeguarding over time.

Figure 1: Mental Health SEA/SH Safeguarding Framework (MHSSF)



7.4 Operationalising the MHSSF

The value of the MHSSF lies not only in its conceptual contribution but also in its practical application across multiple components of mental health systems.

Policy and Legislation

National mental health policies, safeguarding legislation, and professional regulatory frameworks should explicitly integrate safeguarding responsibilities within service delivery. Rather than focusing solely on misconduct investigations, policy should promote preventive governance, leadership accountability, survivor-centred reporting, and organisational learning.

Clinical Practice

Safeguarding should become an integral component of routine clinical care. Mental health practitioners should continuously assess Dynamic Vulnerability, recognise Therapeutic Power Asymmetry, maintain professional boundaries, and utilise reflective supervision throughout clinical practice.

Embedding safeguarding into everyday therapeutic work strengthens patient safety while preserving therapeutic effectiveness.

Professional Education

Safeguarding competence should be embedded throughout undergraduate, postgraduate, internship, residency, and continuing professional education. Training should extend beyond legal obligations to include:

- Professional boundaries;
- Ethical decision-making;
- Management of therapeutic power;
- Trauma-informed care;
- Recognition of grooming behaviours;
- Responding to disclosures; and
- Reflective clinical practice.

Developing safeguarding competence early in professional training strengthens prevention before practitioners enter independent practice.

Organisational Governance

Mental health organisations should regularly evaluate safeguarding maturity by assessing:

- Leadership commitment;
- Governance structures;
- Workforce competence;
- Supervision quality;
- Reporting systems;
- Organisational culture; and
- Continuous improvement mechanisms.

Routine organisational self-assessment enables institutions to identify weaknesses before safeguarding failures occur.

Quality Improvement

The MHSSF positions safeguarding alongside patient safety, clinical governance, and quality improvement as a defining feature of high-quality mental health care. Incorporating safeguarding indicators into accreditation and performance monitoring promotes organisational accountability and continuous improvement.

7.5 Relevance for Kenya and Global Mental Health

Although informed by international literature, the MHSSF is particularly relevant to low- and middle-income countries undergoing mental health reform.

Kenya has strengthened its commitment to mental health through legislative and policy initiatives to improve access, protect human rights, and promote community-based services. Existing legal instruments, including the Constitution of Kenya, the Mental Health (Amendment) Act, the Health Act, the Sexual Offences Act, the Employment Act, the Occupational Safety and Health Act, and professional ethical codes, provide important

safeguards against exploitation and abuse. However, these protections remain dispersed across multiple sectors and have not been integrated into a mental health-specific safeguarding framework.

The MHSSF bridges this gap by linking legal obligations, professional ethics, organisational governance, and therapeutic practice into a coherent safeguarding model. The framework offers practical guidance for policymakers, professional councils, universities, healthcare institutions, and regulators seeking to strengthen safeguarding within mental health systems.

Beyond Kenya, the framework is applicable to community mental health services, psychiatric hospitals, humanitarian mental health and psychosocial support programmes, refugee settings, forensic mental health services, substance use treatment programmes, disability services, and digital mental health platforms. Although implementation should be adapted to local legal and cultural contexts, the underlying principles of prevention, accountability, therapeutic integrity, and organisational learning are broadly transferable.

7.6 Future Directions

The MHSSF is proposed as a conceptual framework requiring empirical evaluation across diverse mental health settings.

Future research should focus on four priorities:

First, empirical studies are needed to determine the prevalence, characteristics, and organisational determinants of SEA/SH within mental health services across different countries and service models.

Second, validated instruments should be developed to measure Dynamic Vulnerability, Therapeutic Power Asymmetry, Boundary Erosion, and Organisational Safeguarding Capacity. Such measures would facilitate benchmarking, organisational assessment, and evaluation of safeguarding interventions.

Third, implementation research should examine how the MHSSF performs across public, private, community-based, humanitarian, forensic, and digital mental health services.

Finally, safeguarding indicators should be incorporated into accreditation standards, workforce competency frameworks, national mental health information systems, and quality improvement programmes to ensure safeguarding is monitored alongside patient safety and service quality.

Advancing Mental Health Safeguarding as a recognised interdisciplinary field will require collaboration among researchers, clinicians, educators, service users, professional associations, and policymakers.

7.7 Conceptual Synthesis

The MHSSF represents the principal contribution of this review by integrating clinical ethics, safeguarding science, trauma-informed care, organisational governance, implementation science, and human rights into a single conceptual model.

Within this framework:

- **Dynamic Vulnerability** explains changing safeguarding needs.
- **Therapeutic Power Asymmetry** explains how professional authority simultaneously facilitates healing and creates safeguarding responsibilities.

- **Boundary Erosion** describes the progressive mechanism through which safeguarding failures develop.
- **Organisational Safeguarding Capacity** determines whether organisations prevent, detect, respond to, and learn from safeguarding risks.

Collectively, these constructs establish Mental Health Safeguarding as an emerging interdisciplinary field extending beyond the prevention of misconduct to encompass prevention, ethical practice, organisational accountability, and continuous improvement.

The MHSSF therefore offers both a conceptual and practical foundation for strengthening safe, ethical, rights-based, and accountable mental health systems in Kenya and internationally.

8. Discussion

This review advances the argument that safeguarding should be recognised as a fundamental dimension of quality mental health care rather than as a narrow response to professional misconduct. Existing literature has traditionally examined professional ethics, therapeutic boundaries, patient safety, workplace harassment, trauma-informed care, and organisational governance as largely independent areas of inquiry. While these bodies of knowledge have significantly informed clinical practice, they provide limited explanation of how safeguarding risks emerge through the interaction of individual, relational, organisational, and systemic factors. The Mental Health SEA/SH Safeguarding Framework (MHSSF) addresses this gap by integrating these perspectives into a coherent systems model.

A major contribution of the framework is the reconceptualisation of vulnerability. Conventional safeguarding approaches often categorise individuals as either vulnerable or not vulnerable based on diagnosis, age, disability, or legal status. Such classifications help identify populations in need of protection but may unintentionally reinforce paternalistic assumptions and overlook the dynamic nature of recovery. The concept of Dynamic Vulnerability proposed in this review recognises that susceptibility to exploitation changes over time according to clinical condition, trauma history, social support, organisational context, and therapeutic engagement. Consequently, safeguarding should be viewed as a continuous and responsive process rather than a static assessment conducted at a single point in time.

The framework also contributes to safeguarding scholarship by explicitly recognising Therapeutic Power Asymmetry as an inherent characteristic of mental health care. Unlike other healthcare encounters, mental health treatment depends on emotional disclosure, trust, confidentiality, and prolonged interpersonal relationships. These characteristics create opportunities for healing but also generate ethical responsibilities arising from unequal authority and influence. Rather than viewing professional authority as problematic, the MHSSF argues that therapeutic power should be exercised transparently, supported through reflective supervision, and regulated by robust organisational safeguards. This perspective shifts attention from individual blame towards responsible stewardship of professional authority.

Another important contribution is introducing Boundary Erosion as the mechanism through which safeguarding failures often develop. Evidence suggests that serious ethical violations are rarely spontaneous events but often evolve through incremental departures from accepted professional standards. By identifying early indicators such as excessive personal disclosure, inappropriate communication, secrecy, preferential treatment, and emotional dependency, the framework emphasises prevention rather than retrospective investigation. This perspective

aligns safeguarding with patient safety approaches that prioritise early risk detection and organisational learning over reactive disciplinary responses.

The MHSSF further expands safeguarding by introducing Organisational Safeguarding Capacity, highlighting the responsibility of healthcare organisations to create environments that actively prevent abuse. Safeguarding failures frequently reflect weaknesses in governance, leadership, supervision, reporting systems, workforce competence, and organisational culture rather than solely the misconduct of individual practitioners. Embedding safeguarding within clinical governance, quality improvement, accreditation, and workforce development enables organisations to strengthen prevention while promoting transparency and accountability. In this regard, safeguarding becomes an organisational capability that requires sustained investment rather than a procedural requirement activated only after incidents occur.

The framework is particularly relevant for Kenya and other low- and middle-income countries, where mental health systems continue to expand amid resource constraints, workforce shortages, and evolving regulatory environments. Kenya has made significant progress in strengthening mental health policy and legislative protections through constitutional guarantees, mental health reforms, and professional regulatory mechanisms. However, safeguarding responsibilities remain dispersed across multiple legal and institutional frameworks. The MHSSF provides an integrated conceptual foundation capable of harmonising these existing mechanisms while supporting implementation across diverse mental health settings. Although developed using Kenya as a contextual lens, the framework is equally applicable to global mental health systems because the principles of professional accountability, trauma-informed care, human rights, and organisational learning transcend national contexts.

Several limitations should be acknowledged. As a critical narrative review, the paper synthesises conceptual and multidisciplinary evidence rather than providing a systematic review or meta-analysis of empirical studies. Consequently, the proposed framework remains theoretical and requires empirical validation across different mental health settings. Future research should develop validated measurement instruments for the four constructs, examine organisational determinants of safeguarding effectiveness, and evaluate implementation strategies across public, private, humanitarian, digital, and community mental health services. Such research will strengthen the evidence base for Mental Health Safeguarding as an emerging interdisciplinary field.

9. Conclusion

Sexual exploitation, abuse, and harassment remain among the most serious threats to patient safety, professional integrity, and human rights within mental health care. Although safeguarding has become firmly established across humanitarian, education, disability, and healthcare sectors, its application within mental health has remained comparatively fragmented, with professional ethics, therapeutic boundaries, patient safety, trauma-informed care, and organisational governance often addressed in isolation. This fragmentation has limited the development of comprehensive safeguarding approaches capable of addressing the unique relational dynamics that characterise mental health practice.

This review argues that safeguarding should be recognised as a core component of quality mental health care rather than merely a response to professional misconduct. Therapeutic relationships are inherently characterised by unequal distributions of authority, specialised knowledge, emotional influence, and institutional legitimacy. These asymmetries are essential for effective treatment but also create ethical obligations that require the continuous protection of service users. Safeguarding therefore cannot be understood solely as preventing individual misconduct; it must encompass the governance, organisational culture, professional education, and leadership that shape everyday clinical practice.

The principal contribution of this paper is the development of the Mental Health SEA/SH Safeguarding Framework (MHSSF), which integrates safeguarding science, professional ethics, trauma-informed care, organisational governance, and human rights into a unified conceptual model. The framework advances four complementary constructs: Dynamic Vulnerability, Therapeutic Power Asymmetry, Boundary Erosion, and Organisational Safeguarding Capacity. Together, these constructs explain how safeguarding risks emerge, how they can be identified before harm occurs, and how organisations can strengthen prevention through systemic rather than purely reactive approaches.

The concept of Dynamic Vulnerability challenges static assumptions regarding vulnerability by recognising that safeguarding needs fluctuate throughout the therapeutic journey. Therapeutic Power Asymmetry explains why professional authority, although necessary for healing, requires careful ethical management and accountability. Boundary Erosion illustrates how serious safeguarding failures frequently evolve through incremental departures from professional standards, highlighting opportunities for early intervention. Organisational Safeguarding Capacity demonstrates that safeguarding outcomes depend not only on individual practitioners but also on leadership, governance, workforce competence, supervision, reporting systems, and organisational culture.

Collectively, these concepts reposition safeguarding as a continuous organisational responsibility embedded within clinical governance, quality improvement, and professional education. The framework therefore moves beyond compliance-based approaches towards a proactive model centred on prevention, transparency, accountability, survivor-centred practice, and continuous organisational learning.

The MHSSF also establishes Mental Health Safeguarding as an emerging interdisciplinary field that bridges psychology, psychiatry, social work, nursing, counselling, public health, implementation science, organisational behaviour, law, and human rights. By integrating knowledge across these disciplines, the framework provides a common language for researchers, educators, practitioners, policymakers, and regulators seeking to strengthen safeguarding within increasingly complex mental health systems.

For Kenya, the framework offers a practical foundation for integrating safeguarding into mental health legislation, professional regulation, workforce education, accreditation standards, and organisational governance. Internationally, the MHSSF provides a flexible model that can be adapted to different legal, cultural, and institutional contexts while maintaining the universal principles of dignity, safety, accountability, equity, and respect for human rights.

Future research should focus on empirically validating the framework, developing safeguarding indicators for mental health services, evaluating implementation strategies, and examining safeguarding across diverse service environments, including community mental

health, psychiatric hospitals, humanitarian settings, digital mental health platforms, and forensic services. Strengthening the evidence base will facilitate refinement of the framework and support its translation into policy and practice.

Ultimately, safeguarding should no longer be regarded as a peripheral administrative obligation but as a defining characteristic of ethical, person-centred, trauma-informed, and rights-based mental health care. Organisations that successfully integrate safeguarding into governance, leadership, professional practice, and continuous learning will be better positioned to protect service users, strengthen public trust, support workforce integrity, and improve the quality and safety of mental health services. The MHSSF provides a conceptual roadmap for achieving these objectives and offers a foundation upon which future scholarship, policy development, and implementation efforts can build.

References

- American Psychological Association. (2017). *Ethical principles of psychologists and code of conduct* (2002, amended effective January 1, 2017). <https://www.apa.org/ethics/code>
- Bordin, E. S. (1979). The generalizability of the psychoanalytic concept of the working alliance. *Psychotherapy: Theory, Research & Practice*, 16(3), 252–260. <https://doi.org/10.1037/h0085885>
- Edmondson, A. C. (1999). Psychological safety and learning behavior in work teams. *Administrative Science Quarterly*, 44(2), 350–383. <https://doi.org/10.2307/2666999>
- Fineman, M. A. (2008). The vulnerable subject: Anchoring equality in the human condition. *Yale Journal of Law & Feminism*, 20(1), 1–23.
- Gabbard, G. O. (2016). *Boundaries and boundary violations in psychoanalysis* (2nd ed.). American Psychiatric Publishing.
- Gelso, C. J., & Hayes, J. A. (2007). *Countertransference and the therapist's inner experience: Perils and possibilities*. Lawrence Erlbaum Associates.
- Grant, M. J., & Booth, A. (2009). A typology of reviews: An analysis of 14 review types and associated methodologies. *Health Information & Libraries Journal*, 26(2), 91–108. <https://doi.org/10.1111/j.1471-1842.2009.00848.x>
- Greenhalgh, T., Thorne, S., & Malterud, K. (2018). Time to challenge the spurious hierarchy of systematic over narrative reviews? *European Journal of Clinical Investigation*, 48(6), e12931. <https://doi.org/10.1111/eci.12931>
- Gutheil, T. G., & Gabbard, G. O. (1993). The concept of boundaries in clinical practice: Theoretical and risk-management dimensions. *American Journal of Psychiatry*, 150(2), 188–196. <https://doi.org/10.1176/ajp.150.2.188>
- Herman, J. L. (1992). *Trauma and recovery: The aftermath of violence—from domestic abuse to political terror*. Basic Books.
- Inter-Agency Standing Committee. (2015). *Guidelines for integrating gender-based violence interventions in humanitarian action: Reducing risk, promoting resilience and aiding recovery*. IASC. <https://interagencystandingcommittee.org>
- Levine, C., Faden, R., Grady, C., Hammerschmidt, D., Eckenwiler, L., & Sugarman, J. (2004). The limitations of "vulnerability" as a protection for human research participants.

- American Journal of Bioethics*, 4(3), 44–49.
<https://doi.org/10.1080/15265160490497083>
- Luna, F. (2009). Elucidating the concept of vulnerability: Layers not labels. *International Journal of Feminist Approaches to Bioethics*, 2(1), 121–139.
<https://doi.org/10.3138/ijfab.2.1.121>
- Norcross, J. C., & Lambert, M. J. (2018). Psychotherapy relationships that work III. *Psychotherapy*, 55(4), 303–315. <https://doi.org/10.1037/pst0000193>
- Reason, J. (2000). Human error: Models and management. *BMJ*, 320(7237), 768–770.
<https://doi.org/10.1136/bmj.320.7237.768>
- Vincent, C. (2010). *Patient safety* (2nd ed.). Wiley-Blackwell.
- Wampold, B. E., & Imel, Z. E. (2015). *The great psychotherapy debate: The evidence for what makes psychotherapy work* (2nd ed.). Routledge.
- World Bank. (2020). *Preventing sexual exploitation and abuse and sexual harassment (SEA/SH) in investment project financing involving major civil works*. World Bank.
<https://www.worldbank.org>
- World Health Organization. (2013). *Responding to intimate partner violence and sexual violence against women: WHO clinical and policy guidelines*. World Health Organization. <https://www.who.int/publications/i/item/9789241548595>