

Trust, Tone, Timing, and Tech. Assessing Digital P/CVE Counselling Services and Challenges for Evaluation Practice

Juliane Kanitz | Maximilian Campos Ruf |
Svetla Koynova | David Tschöp | Franziska Heil

PrEval Studie 6/2025

TRUST, TONE, TIMING, AND TECH. ASSESSING DIGITAL P/CVE COUNSELLING SERVICES AND CHALLENGES FOR EVALU- ATION PRACTICE

Juliane Kanitz | Maximilian Campos Ruf |
Svetla Koynova | David Tschöp | Franziska Heil



Peace Research Institute Frankfurt
Leibniz-Institut für
Friedens- und Konfliktforschung

PREVAL – ZUKUNFTSWERKSTÄTTEN EVALUATION AND QUALITY MANAGEMENT IN EXTREMISM PREVENTION, DEMOCRACY PROMOTION AND CIVIC EDUCATION: ANALYSIS, MONITORING, DIALOGUE

C/O PRIF – PEACE RESEARCH INSTITUTE FRANKFURT
C/O PRIF – LEIBNIZ-INSTITUT FÜR FRIEDENS- UND KONFLIKTFORSCHUNG

Cover:

The cover image was created using artificial intelligence (AI).

The following prompt was used: “Abstract architectural structure (bright, not so dark image)”

Creation and editing: grübelfabrik

Text license:

Creative Commons CC-BY-ND 4.0 (Attribution/No Derivatives/4.0 International).

The images used are subject to their own licenses.



Correspondence to:

Peace Research Institute Frankfurt
Darmstädter Landstrasse 112
D-60598 Frankfurt am Main/Germany
Telephone: +49 69 95 91 04-0
E-Mail: preval@prif.org
<https://www.prif.org> | <https://preval.hsfk.de/en>

DOI: 10.48809/PrEvalStu2506

About the Authors

Juliane Kanitz, Dr., is a scholar of Islamic Studies and a social anthropologist. She heads evaluation research at *i-unite* and develops methods to make impact and learning processes in online prevention visible. Her focus is on practice-oriented, qualitatively grounded evaluation.

Maximilian Campos Ruf is a scholar of Islamic Studies and social scientist with a focus on terrorism and security studies. He is the head of the Research Department at *Violence Prevention Network gGmbH*. His current areas of focus include international cooperation, collaboration between security agencies and civil society, MEL processes, and the study of hybrid radicalisation phenomena.

Svetla Koynova is a political scientist and a sociologist of religion with a focus on intercultural and transcultural management. Together with Maximilian Campos Ruf, she heads the Research Department at *Violence Prevention Network gGmbH*. Her research focuses include monitoring, evaluation, and learning processes, group-focused enmity, and radicalisation.

David Tschöp is a social scientist. From 2022 to 2024 he first worked as a student assistant and then as a research associate in the Research and International departments at *Violence Prevention Network gGmbH*.

Franziska Heil is a pedagogue and criminologist whose work focuses on extremist strategies and prevention at the intersection of research, practice, and public administration. At the *Peace Research Institute Frankfurt (PRIF)*, she works in the project *PrEval – Zukunftswerkstätten* and is a member of the *Research Group on Radicalization, Terrorism and Extremism Prevention*. She also works freelance in youth care and teaches “Extremism” at the *Hessian University of Public Management and Security (HöMS)*, Department of Police.

The Authors According to Their Own Words

Juliane Kanitz | Maximilian Campos Ruf | Svetla Koynova | David Tschöp | Franziska Heil

Which disciplinary perspectives do we hold?

We focus on the phenomenon from anthropology, Islamic studies, political and social sciences perspectives, and most of us work in organisations engaged in pedagogical and social work practice. This interdisciplinary perspective allows us to analyse digital counselling and prevention work in its social, cultural, and technological dimensions. We combine empirical research with practice-oriented evaluation and build bridges between qualitative analysis and institutional programme development.

What is our research interest?

We examine how digital prevention work and online counselling in the field of P/CVE can be evaluated within the context of extremism prevention. At the centre are questions of practical utility, impact, trust, ethics, data protection, and technology. Our aim is to support the development of realistic evaluation designs that combine methodological rigor with practical feasibility while also fostering learning processes for practitioners and organizations.

How does our study contribute to the *PrEval* fields of action?

Our study strengthens *PrEval* by providing empirically grounded insights into digital prevention practice and by developing practice-oriented methodological recommendations for evaluation and quality assurance. It outlines ways to reconcile impact measurement, data protection, and learning processes, thereby contributing to the further development of evidence-based, ethically reflective prevention work.

Die *PrEval* Studies

Evaluation generates knowledge—but it can only be successful if knowledge is incorporated into its design: knowledge about the specifics of the respective field, about suitable methods, and about the dynamics in which measures take effect. At the same time, openness and a willingness to jointly develop processes are required. This is precisely where *PrEval* comes in. We bring together practitioners, academics, administrators, and funding institutions to analyze needs, document experiences, and jointly develop sound, practical approaches to evaluation and quality management. This results in reliable findings that not only improve existing measures but also help to design new formats in a more targeted and effective way. Evaluation is not seen as a control instrument, but as a contribution to a learning and reflective prevention practice. *PrEval* creates spaces for this, provides impetus, and strengthens the dialogue between the groups of actors involved.

As an independent research and transfer project, *PrEval* has been working since 2020 to strengthen quality management and evaluation in the fields of extremism prevention, civic education, and democracy promotion. Our work focuses on analyzing existing capacities and developing and testing evaluation designs and support formats that are based on real needs – in continuous dialogue with funding agencies, practitioners, administrators, academics, and donors. Particular attention is paid to the interfaces between stakeholder groups, responsibilities, program logics, and institutional frameworks. *PrEval* pursues a multi-method, practice-oriented approach to promote knowledge, exchange, and trust among stakeholders, thereby improving the conditions for sustainable evaluation.

The *PrEval Studies* build on the earlier *PrEval* report series. They focus on concrete evaluation experiences from practice—for example, on digital formats for civic education, evaluation in complex counseling settings, or cooperation between civil society organizations and security authorities. The studies show how evaluation processes can be meaningfully designed together, where obstacles exist, and which approaches are feasible in practice. They are aimed at practitioners, academics, funding institutions, and decision-makers, and are intended as practical impetus for a learning prevention landscape, but also as impetus for the academic evaluation debate.

Frankfurt am Main, November 2025

Overview of All Titles in the *PrEval Studies Series*

Entwicklung eines integrierten Modells und Selbsteinschätzungsinstruments für *Digital Citizenship Literacy* by Marcus Kindlinger // Lucy Huschle // Hermann Josef Abs. PrEval Studie 1/2025.

Entwicklung eines KI-Kompetenzprofils aus Perspektive der politischen Bildung by Marcus Kindlinger // Hermann Josef Abs. PrEval Studie 2/2025.

Bedarfsorientierte Evaluation in sicherheitsrelevanten Kooperationskontexten. Konzeption eines Evaluationsdesigns by Juliane Kanitz. PrEval Studie 3/2025.

Wirkungshorizonte von Evaluationen: Zum Uptake von Evaluationen der Extremismusprävention by Marion Lempp. PrEval Studie 4/2025.

Digitalgestützte Evaluation von Digitalprojekten der politischen Bildung: „Ein Klick ist nicht gleichbedeutend mit produktiver Auseinandersetzung“ by Susanne Johansson // Andrea Prytula. PrEval Studie 5/2025.

Trust, Tone, Timing, and Tech. Assessing Digital P/CVE Counselling Services and Challenges for Evaluation Practice by Juliane Kanitz // Maximilian Campos Ruf // Svetla Koynova // David Tschöp // Franziska Heil. PrEval Studie 6/2025.

Klien*innenzentrierte Evaluation in komplexen Beratungssettings by Svetla Koynova // Emma van Heeswijk.

Citizenship Education and the Prevention of Radicalisation and Extremism. An International Comparative Study into the Organisation of Preventing Radicalisation and Extremism within (Citizenship) Education by Myrte van Veldhuizen // Hermann Josef Abs.

Summary

This report investigates how the current state of the art of digital counselling in the field of preventing and countering violent extremism (P/CVE) can be understood and meaningfully evaluated. It draws on an in-depth analysis of German practice while situating findings within broader international developments. Over the past decade, digital counselling has evolved from a marginal add-on to a central pillar of prevention work. What once happened face-to-face now unfolds in chats, video calls, and gaming or social media environments. This expansion has opened new avenues for access and outreach but also exposed deep methodological and ethical challenges for evaluation, professionalisation, and accountability.

The study addresses these challenges through a two-stage qualitative design combining desk research with eight expert interviews and two virtual focus groups. The analysis distinguishes two principal logics of digital counselling: reactive formats, where clients seek help through protected channels such as encrypted chats or video sessions; and outreach formats, where practitioners proactively engage potential clients in public or semi-public digital spaces such as *Instagram*, *Discord*, or gaming platforms. Both logics rely on trust-building in often unstable digital environments, yet they operate with radically different rhythms, infrastructures, and evidentiary possibilities. Acting as a bridge between the two, referral-based models focus on triage, trust, and redirection to suitable services.

Across all types, technology is not merely a tool but a co-producer of practice. Platform architecture, algorithmic visibility, and interface design shape what can be said, recorded, or measured. This entanglement defines the main evaluative dilemma: the very conditions that enable digital access, levels of anonymity, ephemerality, and decentralization, also limit a systematic assessment of outcomes. Practitioners thus navigate ethical restraint and the demand for evidence, improvising with screenshots, internal logs, and qualitative reflections to make their impact visible.

The findings show that evaluation in digital P/CVE counselling must evolve from a compliance exercise to a learning-oriented, adaptive process. Standard indicators such as reach or session counts reveal little about relational depth or behavioural change. Instead, mixed-method designs that combine quantitative monitoring with qualitative case analysis are required. Evaluations should capture not only whether something works, but how, for whom, and under which platform conditions change becomes possible. Process-oriented and formative approaches, joint reflection sessions, peer reviews and practitioner diaries prove more useful for professional development than one-off outcome measurements.

A persistent structural gap remains between funders' accountability frameworks and practitioners' realities. Funders often expect linear proof of impact, while frontline workers ex-

perience fragmented, low-visibility interactions that defy such metrics. Building a shared evaluation language and establishing long-term partnerships between funders, practitioners, and researchers emerge as central recommendations. Evaluation must become a collaborative practice embedded in daily work, not an external audit applied after the fact.

Finally, the report briefly explores the emerging role of artificial intelligence. AI-assisted tools could potentially support documentation, early risk detection, and analysis of large chat datasets. Simultaneously, they raise substantial ethical and practical concerns, data protection, bias, and the loss of relational sensitivity. AI should assist, not replace, human judgment, and its integration requires transparency, co-design with practitioners, and continuous oversight.

In conclusion, evaluating digital counselling in P/CVE is both necessary and possible, but only if methods adapt to the specific logics of digital practice. Effective evaluation in this field is not about counting interventions but about understanding relationships, reflection, and adaptation within volatile digital environments. When treated as a means of collective learning rather than bureaucratic control, evaluation becomes part of the professional craft itself, helping to make the invisible work of digital prevention visible, credible, and sustainable.

Contents

Summary	9
1. Framing the Challenges of Digital P/CVE Counselling	13
2. Methodology: Investigating Digital Counselling in P/CVE	14
2.1 Desk Research	15
2.2 Expert Interviews	15
2.3 Focus Groups	16
2.4 Analytical Approach	16
3. Structuring Digital Interventions Leading to Counselling in P/CVE Work	16
3.1 Towards a Definition of Digital Counselling in P/CVE	16
3.2 Types of Digital Interventions Leading to Counselling in P/CVE: Goals and Formats	19
3.3 How Technology Shapes Counselling	21
3.4 Platform Logic and Its Impact on Counselling Practice	22
3.5 Counselling Diverse Audiences, Diverse Needs: Mapping Digital Counselling Users	24
4. When Measurement Meets Reality: Rethinking Evaluation in Digital Interventions and Counselling in P/CVE	26
4.1 Assessing Levels of Professionalism: Developing Stance and Standards	27
4.2 Reaching Users Online: Why Evaluation Must Adapt	28

4.3 Evaluating Impact: Between Expectations and Evidence	31
4.4 Bound by the Platform: How Tech Shapes Impact	35
4.5 Adapting on the Fly: Digital Counselling Without a Roadmap	39
4.6 Too Much, but Never Enough: Data Limits in Digital Counselling	41
4.7 Doing More with Less: Practical Approaches to Evaluation	44
4.8 Lost in Translation: Funders, Researchers, and Practitioners	46
5. Looking towards the Future: From Hype to Help? The Potential of AI in Digital Counselling	48
6. What We've Learned – And Where to Go Next	50
Focus Groups and Expert Interviews	53
Literature	55
Appendix	57

1. Framing the Challenges of Digital P/CVE Counselling

Evaluating digital counselling services in preventing and countering (violent) extremism (P/CVE) is anything but straightforward. On the one hand, there is pressure to show measurable results. On the other hand, the work often takes place in fleeting, anonymous online spaces that resist easy analysis. Indeed, in many instances, online outreach and counselling may not lead to prolonged counselling relationships but may instead occur as singular interactions. In addition, counsellors and other service providers operate across different types of platforms (e.g., *TikTok*, *YouTube*), many of which were never intended to foster the levels of deep conversation necessary for counselling services. Relationships matter in counselling settings, but they are often more difficult to establish in digital environments. At the same time, funders seek demonstrable effectiveness of such interventions, but the real value lies in slow, trust-based work that is often hard to quantify. Evaluating digital counselling in P/CVE is thus complex, due to its anonymous, evolving, and transient nature. Therefore, this report specifically aims to move beyond mere problem description and towards identifying practical evaluation methods and metrics that effectively capture impact in spite of these challenges, in the hope of providing readers with concrete guidance informed by practice.

The two fields, digital counselling in P/CVE and its evaluation, are still trying to find their footing, both in terms of practice and regarding relevant research. New methods, such as blended evaluation models, tools that respect user privacy, and approaches that involve practitioners in the process, are starting to bridge the gap between ethical practice and meaningful learning. Increasingly, many in the field now see evaluation not as a burden, but as a way to reflect, adapt, and grow. In order to move forward, evaluation needs to shift away from checklists and become a flexible means, tailored to the fast-changing, sensitive world of digital prevention and counselling. This requires joint efforts from funders, researchers, and practitioners alike, along with a rethinking of what 'evidence' means in this space.

The real question is not just whether something works, but how, why, for whom, and in what contexts it works. In countries such as Germany, what started twenty years ago as traditional face-to-face support for radicalised people has evolved, especially during the COVID-19 pandemic, into hybrid and digital approaches, comprising text messaging, live chats, phone calls, email, and video calls. Each of those communication pathways comes with its own respective challenges regarding data and client safety, effectiveness, and other factors. The internet is no longer simply a tool; it is where life happens. As a result, P/CVE practitioners today should be familiar with the basic tools and considerations necessary for implementing digital or hybrid counselling approaches in a post-digital society.

This broader reach is both a gift and a challenge. Counsellors now work across chat apps and other platforms, with some also adopting AI tools, often wondering how to demonstrate the impact of what they do when clients stay anonymous and conversations are scattered and brief. That is where this *PrEval Study* comes in: assessing the state of the art of digital counselling in P/CVE today, learning from international pioneers, and asking what might be getting in the way, along with exploring exciting new possibilities for evaluating online impact in P/CVE contexts.

Three guiding questions frame this report:

- **What psychosocial approaches and logics** of digital counselling currently shape secondary and tertiary extremism prevention, and how do platform dynamics influence these formats?
- **What methodological challenges** arise when evaluating the impact of digital counselling, particularly regarding anonymity, data limitations, and rapidly changing environments?
- Lastly, considering recent rapid developments around artificial intelligence, the report includes an excursion to briefly address some key issues around the question of: **How can emerging technologies, especially AI**, help to overcome these evaluative challenges, and what ethical and practical limitations need to be considered?

Chapter 2 begins by defining and classifying digital counselling in P/CVE. Chapter 3 then proceeds to unpack its evaluative pitfalls and pathways, followed by a look ahead towards AI-enabled futures that may shape the field. The report therefore provides both a snapshot of Germany's current digital P/CVE landscape and a compass for anyone determined to make impact visible in the virtual realm.

2. Methodology: Investigating Digital Counselling in P/CVE

This report is based on a two-stage qualitative design combining (1) desk research with (2) expert interviews and (3) focus groups. The methodological approach builds on the foundations developed in the *Evaluation streetwork@online* study (Kanitz et al. 2021), the first systematic evaluation of online streetwork in extremism prevention in Germany. That study applied a qualitative, grounded-theory design combining document analysis, expert interviews, and participatory observation to explore how digital outreach unfolds on social media platforms. It emphasized professionalization, ethical boundaries, and the learning dynamics within online prevention practice, framing evaluation as a formative and reflective process rather than a measurement exercise. These principles inform the present study's approach to analysing and evaluating digital counselling practices, while adapting

procedures to the specific characteristics of digital counselling in secondary and tertiary prevention.

2.1 Desk Research

An initial desk research phase was conducted to map existing digital counselling services in the German P/CVE field and identify international reference projects. Search terms included combinations such as *digital counselling extremism*, *online outreach prevention*, *AI AND P/CVE*, and *messenger-based interventions*. Sources included academic literature, grey publications, project websites, and internal strategy papers. Selection criteria prioritised sources demonstrating:

- Active P/CVE focus (excluding general mental health hotlines)
- Clear use of digital channels (especially chat/messenger-based)
- Evidence of evaluation or monitoring and documentation efforts
- Institutional transparency and identifiable contact structures

Cases were excluded if they lacked relevance for the German-speaking context or referred to one-off campaigns without sustained structures. In sum, nine relevant practices were identified for Germany, substituted by another six international reference projects, all of which were active at the time of data collection in August 2024.

2.2 Expert Interviews

To deepen insights into practical challenges and innovations, 8 semi-structured interviews were conducted with practitioners from two subfields:

- Classic digital counselling services, often integrated into existing structures and operating via encrypted chats, email, or video sessions.
- Outreach-based digital formats, which proactively engage target groups on platforms like *Instagram*, *TikTok*, or *Discord*.

Participants were selected for their hands-on experience and institutional diversity (NGOs, umbrella organisations, tech-supported pilot projects) and based on the results of the desk research. Interviews were conducted online, in German and English, recorded with consent, and transcribed verbatim using *MAXQDA*. Interview questions focused on ethical dilemmas, platform logics, evaluation tools, and practitioner learning cultures.

2.3 Focus Groups

Two virtual focus groups were held to validate interview findings and expand on key themes:

- One group brought together six **practitioners from classic counselling settings**, discussing depth and continuity pertaining to client relationships in digital formats.
- The other group included five **outreach workers**, focusing on issues such as identity fluidity, algorithmic visibility, and fragmented interactions.

Each focus group session lasted 90 minutes and was structured around guiding questions informed by prior interview analysis. Notes were taken collaboratively by two facilitators and coded using a thematic coding framework.

2.4 Analytical Approach

Data analysis followed a **thematic synthesis approach**, drawing on *Grounded Theory* elements (Glaser/Strauss 1967). Codes were developed inductively, clustered iteratively, and discussed in interdisciplinary interpretation teams. Emerging themes were continuously validated against prior findings from Kanitz et al. (2021) and Koynova (2025), especially regarding evaluation in dynamic, trust-sensitive cooperation contexts.

The final report synthesises findings from the desk research, interviews, and focus groups into practical guidance for evaluating digital counselling within the complexities of P/CVE. Results of the desk research were primarily used to inform the analysis of the existing digital interventions and counselling landscape in chapter 3. The empirical research conducted through focus groups and expert interviews form the primary basis for the exploration of relevant challenges for practice and evaluation in this field, as set out in chapter 4.

3. Structuring Digital Interventions Leading to Counselling in P/CVE Work

3.1 Towards a Definition of Digital Counselling in P/CVE

This section provides both a theoretical overview of what digital counselling encompasses, as well as a presentation and discussion of different approaches in the context of P/CVE.

Counselling is a key concept across social pedagogy, social work, and P/CVE. According to the *Organisation for Security and Co-operation in Europe* (OSCE), P/CVE refers to

non-coercive activities aimed at preventing or mitigating violent extremism by addressing its underlying drivers. Within this context, it is important to analytically distinguish between prevention and counselling (OSCE 2019).

Prevention focuses on the 'who,' 'when,' and 'why' of interventions and is traditionally segmented into three levels, building on Caplan (1964):

- **Primary prevention** targets the general population and aims to strengthen democratic values, civic resilience, and social justice;
- **Secondary prevention** addresses at-risk individuals, including those showing early signs of radicalisation; and,
- **Tertiary prevention** focuses on already radicalised individuals, including those involved in extremist or terrorist activities (e.g., Molenkamp et al. 2018: 2).

Counselling, as a method commonly used in social work or therapy, focuses on the 'how'. In the context of P/CVE, counselling is most commonly applied throughout the secondary and tertiary levels, due to their narrower target group definition. It involves structured dialogue processes to support individuals in navigating personal, emotional, social, and sometimes psychological challenges, ideally preventing deeper involvement in extremism (Leong 2008: 119). In P/CVE, counselling is often an essential, though not exclusive, component (Baaken et al. 2018: 4–5), especially within secondary and tertiary prevention contexts.

Despite its growing importance, digital counselling remains difficult to define precisely. Key tensions remain in determining where digital counselling begins and ends, particularly given the blurred boundaries between digital and in-person (i.e., non-technologically mediated) communication. Technical, methodological, temporal, and relational factors must all be considered. Common definitions include a vast array of measures that can constitute digital counselling or digital therapy, with Gehrman describing it as 'computer-supported, media-mediated, and interactive' (2010: 106)' and Gheorghe et al. summarising digital work in the context of psychotherapy as 'rely[ing] on electronic or internet-based forms of communication to deliver mental health care synchronously (i.e., real-time) or asynchronously (i.e., not real-time)' (2024: 717), which might include videoconferences as well as instant messaging, among others. While chatbots are still experimental in most counselling contexts, the current AI boom suggests that more widespread implementation of such systems in counselling and intervention settings might only be a matter of time, regardless of how ready for professional and ethical use they might be at that stage.

The *German Society for Psychosocial Online Counselling* (DGOB) offers a more integrative approach, viewing digital counselling as both an alternative access point and a mechanism for the transformation of interaction, from physical to telemediated, or vice-versa. Central

to this is the notion of physical distance, which might reshape counselling processes and underlying dynamics (Thiery 2022: 4–7). Highlighting the key differences between mediated and in-person counselling, the DGOB emphasises the fact that digital counselling is not an independent *method*, but rather an alternative *setting* for counselling (Thiery 2022: 7). As such, it offers alternative access points to a counselling process, in such cases where physical interaction is not possible: ‘Therefore, the boundaries of one setting, prove to be the opportunities of the other – and vice-versa’ (Thiery 2022: 7). This way of thinking about digital counselling ensures that this new and necessary setting of delivering psychosocial interventions can confidently hold its ground as a professional method, appropriately recognising the deep impact of digital and mediated forms of communication on people’s lives today, as long as it remains rooted in professional and ethical practice.

That being said, digital counselling provides both opportunities and risks. It can promote anonymity and accessibility, for example by reaching underserved populations in rural areas or marginalised groups. It can reduce logistical barriers such as travel or office space costs. At the same time, it introduces a new level of data privacy and security challenges as well as purely practical ones, which all need to be tackled by practitioners. Practical challenges include seemingly simple tasks such as the establishment of a stable and secure digital platform, while ensuring practitioners are able to use them properly. Additionally, and especially in potentially violence-adjacent contexts such as P/CVE work, client safety is a real concern. This includes clients’ ability to be in a safe physical environment during a digital counselling process, or the establishment of ‘first aid kits’ for crisis situations, including emergency contacts and crisis service providers (Van de Donk/Ruf 2022: 4).

Moreover, it is increasingly unhelpful to draw a binary line between online and offline worlds. Many services now operate across this continuum, blending digital initiation or service delivery with in-person follow-up, or the other way around. Services may also be delivered solely in digital contexts, without an in-person component. Consequently, this study adopts a spectrum-based model, focusing on services where the initial contact, major parts of the counselling process, or the entire service delivery occur online. This hybrid nature requires flexible yet rigorous definitions, grounded in both theory and practice.

Our following working definition clarifies the meaning of digital counselling within the context of preventing and countering violent extremism (P/CVE) and can provide orientation for future practical and evaluation development purposes:

Digital counselling in P/CVE is a multifaceted, evolving practice that encompasses formats of structured, trust-based dialogue as well as short-term interventions through (a combination of) diverse technological media. It aims to initiate individual reflection processes and positive change by addressing personal, psychological, and social challenges related to involvement in extremism through various

forms of synchronous and asynchronous communication, including anonymous messaging, video interactions, and potentially AI-supported tools.

These aspects will be elaborated further in subsequent sections, starting with a typology of existing (interventions leading to) counselling approaches in practice.

3.2 Types of Digital Interventions Leading to Counselling in P/CVE: Goals and Formats

Digital counselling in P/CVE can take many forms, depending on both its goals and operational modalities. This section proposes a typology based on counselling logic, distinguishing between direct, referral-based, and hybrid formats. These categories help to clarify what type of interaction is taking place, what kind of relationship is being fostered, and what expectations can reasonably be attached to different counselling configurations.

1. Classic Counselling Models. Classic models of digital counselling aim to replicate in-person counselling dynamics in a digital setting. These formats prioritise relationship-building, regular sessions, and case continuity. They may involve scheduled video or phone calls, or a structured exchange via secure chat or email. Even when organised asynchronously, these formats typically follow clear case protocols and seek to establish a therapeutic alliance (Thiery 2018). In most instances, these approaches are set up similarly to, or even as part of, in-person counselling centres, offering support to persons reaching out on their own, rather than employing outreach methods.

2. Outreach Interventions and Counselling. Digital outreach formats take a different approach. Here, counsellors enter digital spaces in order to proactively contact individuals who may display signs of vulnerability or risk related to radicalisation pathways – but who have not yet expressed a need for support themselves. These environments include message boards and forums (e.g., *Reddit*), comment sections (e.g., on *Twitch* or *YouTube*), gaming-related environments (e.g., *Roblox*, or *Discord*), or social media platforms such as *TikTok* or *X*. Interactions in such settings may be short and discontinuous, but can still have significant impact, for example by sparking reflection or by providing opportunities to de-escalate during acute emotional crises, or to off-ramp persons into longer-term counselling settings. This has been confirmed by practitioners, by previous research (e.g., Kanitz et al. 2021) as well as this report's interviews:

'And it turned out that, for me, but also my colleagues, quick comments, phrases made in passing, without a direct intention for them to be impactful, had a big impact on the person opposite me.' (FGG1, Pos. 14)

It is worth noting that many existing practices of this type explicitly target gaming-adjacent platforms, spaces that seem to offer ample opportunities to contact a wide variety of demographics. In Germany, such services are offered by, for example, *streetwork@online* (AVP e.V.; <https://streetwork.online>), a project that combines direct engagement on *TikTok*, *Instagram*, and *Discord* with offline follow-up and links to broader prevention campaigns.

3. Referral-Based Interventions and Counselling. Some digital P/CVE initiatives do not seek to provide counselling services themselves, but act as intermediaries instead. They aim to build initial trust, offer a safe first point of contact, and then redirect individuals to services appropriate to their respective challenges and/or locations. This logic is common when projects are limited in scope or capacity, or when certain legal or psychological complexities exceed the project's remit. Referral is often organised through a central network spanning different locally and/or thematically situated organisations (OSCE 2019:9). The largest P/CVE-focused examples of this in Germany are the *Federal Agency for Migration and Refugees (BAMF) Advice Centre on Radicalisation* (<https://www.beratungsstelle-radikalisierung.de>), or the *Beratungskompass Verschwörungsdenken* (<https://beratungskompass-verschwoerungsdenken.de/>).

4. Mixed and Hybrid Forms. Many services combine elements from these different logics. For instance, dedicated outreach teams may initiate contact in the digital realm and then hand over to classic counselling teams within the same organisation or refer to adjacent service providers. Some platforms integrate live chats with asynchronous follow-ups, or offer options for either direct help or referral, depending on user preference. Other approaches explore broader intervention options in specific subcultural spaces, paired with elements of online outreach or online-offline referral. The specificities of these hybrid formats often emerge and are adjusted over time in response to fluctuating platform constraints and user needs. One such example is the gaming-focused approach by *Amadeu Antonio Foundation* in Germany, *Good Gaming – Well Played Democracy* (<https://www.amadeu-antonio-stiftung.de/good-gaming-well-played-democracy/digital-streetwork/>).

This typology reveals the diversity of existing digital counselling and counselling-related interventions in current P/CVE practice. It helps evaluators and funders understand the specific affordances and limitations of each approach. It also shows that counselling logic is not merely a technical issue but rather one that fundamentally shapes the kind of support that can be offered, as well as the resulting outcomes that can be expected.

3.3 How Technology Shapes Counselling

While the typology of counselling logics provides a conceptual understanding of different approaches to digital counselling, the actual configuration of services is equally shaped by the technological medium employed. Different formats, such as live chats, email, phone, video, and multi-feature platforms, offer specific affordances and limitations, each influencing the character and depth of interactions.

- 1. Synchronicity: Synchronous vs. Asynchronous.** Digital counselling may unfold in real-time (synchronous) or across temporal delays (asynchronous). Synchronous formats, such as live chats, calls, or video sessions, allow for more immediate emotional regulation, effective crisis management, and relational feedback. Asynchronous communication, such as email or message boards, may allow for greater reflection, linguistic precision, and accessibility across time zones, while also posing certain risks for clients during crises. Both formats have merits and disadvantages that need to be weighed carefully, and that sometimes need to be combined in flexible sequences.
- 2. Text-Based vs. Audio, Visual, or Audio-Visual Media.** Text remains a common medium for digital counselling in P/CVE. Its perceived anonymity, ease of use, and documentation potential make it well-suited for low-threshold entry. However, it can limit emotional nuance and depth. Call or video formats introduce additional sensory channels, voice, facial expressions, and gestures, enhancing emotional attunement but also requiring a higher degree of technological reliability and user confidence. At the same time, audio-visual media require certain conditions to ensure that clients are not harmed in the process, including the ability to conduct counselling sessions from within a physical safe space. Projects must carefully weigh the affordances and risks of each medium, especially when working with particularly sensitive target groups, such as minors or traumatised individuals.
- 3. Interface and Platform Design.** The counselling experience is deeply affected by the interface through which communication occurs. Some services use purpose-built counselling platforms with secure login areas, data-secure chat modules, video-call servers, and note-taking tools. Others rely on free public messenger apps like WhatsApp, Discord, or Telegram, balancing accessibility with limited control over data privacy, interface updates, and user identity. Some platforms restrict media types (e.g., no video or screenshot functions), directly shaping how counselling can be implemented.
- 4. Multi-Channel and Multi-Modal Approaches.** An increasing number of services adopt multi-channel models, combining several media forms depending on user preferences and case progression. A common trajectory might involve first

contact via text, chat, or a brief phone call, followed by a switch to more stable and long-term email-based counselling or video consultation. Some projects offer platform-switching based on risk level or availability. Others coordinate across internal teams, matching users to counsellors depending on the preferred medium or need for intervention depth.

Each technological approach comes with its own technical, professional, and ethical constraints. Low-bandwidth environments, shared devices, and interface friction may discourage use. Conversely, some projects explore innovations such as embedded risk assessment tools, AI-assisted triage, or data visualisation dashboards to monitor communication quality (Khairat et al. 2018; Limbic AI 2025; Ni/Jia 2025). These tools raise important ethical questions, but they also offer new avenues for structured case management and evaluation.

Importantly, technology is not a neutral vessel for counselling; it co-produces the kind of counselling that is possible. Recognising the interplay between medium and counselling logic enables more realistic, context-sensitive service design and evaluation. The next section will examine how these dynamics play out across platforms and in relation to specific user groups.

3.4 Platform Logic and Its Impact on Counselling Practice

The choice of platform architecture and affordances shapes not only technical possibilities but also pedagogical strategies and limitations in digital counselling. Across the field, four prototypical platform logics have emerged:

1. **Chat-Based Messenger Platforms.** Chat-based counselling services offer quick, low-threshold access, especially for younger users familiar with messaging apps like WhatsApp, Signal, or Telegram. These platforms are often promoted through hybrid outreach strategies: digital advertisement paired with analogue means such as printed flyers, in-person workshops, and direct referrals via youth workers or social workers who provide trusted contact points. This combination of online and offline outreach helps bridge the gap between digital entry points and real-world support systems and is in line with non-digital approaches to counselling, which strongly rely on referrals by surrounding professional support systems and adjacent fields to gain client access. These services rely primarily on direct messaging tools, allowing for fast-paced, low-threshold interaction. As there is no live moderation on chat-based platforms (aside from delayed and often futile reports that users have to file themselves), counselling on such platforms is primarily about shaping the relationship work with the person seeking support. While *WhatsApp*, *Discord*, or *Signal* moderate at least some public-fac-

ing content reported for abuse, several peer-reviewed studies describe *Telegram* as having 'less strict moderation' (Urman/Katz 2020: 905), leading to a situation in which 'users communicate in ideologically homogenous groups and distribute extremist content without fearing any legal consequences' (Urman/Katz 2020: 908).

Beyond significant ethical risks attached to the use of such messenger apps, their privacy constraints, lack of metadata access, and limited interface control can present significant challenges for evaluation. For instance, counsellors report difficulties in 'documenting interaction flow' or analysing patterns beyond what the platform interface permits.

- 2. Thread-Based and Forum-Inspired Spaces.** Inspired by *Reddit*, classic forums, or *Discord*, these platforms enable asynchronous exchange within public or semi-public channels. Some, like *Discord*, might bridge asynchronous and real-time options, for example by offering the possibility to organise live voice chats. This architecture supports community-based engagement, the normalisation of experience, and peer support, but it also requires consistent moderation. Projects using *Discord*, for example, report tensions navigating the sometimes conflicting priorities between free speech and safety. If employed as part of a P/CVE counselling approach, the pedagogical opportunity here lies in guided co-reflection: trained moderators and/or counsellors can highlight patterns, provide psychoeducational content, and link users to one-on-one support where needed.
- 3. Algorithmically Curated Platforms.** Social media environments such as *Instagram*, *TikTok*, or *YouTube* follow a distinct logic: discoverability through algorithmic amplification. These platforms can facilitate initial contact, content-driven prevention, and awareness-raising, but they are shaped by opaque ranking mechanisms that promote (or demote) certain materials based on previous platform behaviour or (potential) user engagement, with little regard for ethical concerns. Practitioners report experiences of 'shadow banning' or declining reach when addressing sensitive topics. One actor summarised: 'We have to assume that our content is filtered out simply for containing the word "radicalisation"' (FGG2, Pos. 35, 40). As a result, pedagogical strategy must account not only for user needs but also for constantly evolving platform policies and algorithmic biases.
- 4. Hybrid and Purpose-Built Platforms.** Some initiatives operate through custom-built platforms or hybrid systems that integrate public outreach and secure private counselling environments, for example, the German state-funded deradicalisation program *Wegweiser*, the German referral counselling project addressing conspiratorial thinking called *Beratungskompass Verschwörungsdenken*, the US-based NGO project *The Daily Former Discord* by *Life After Hate* (LAH 2024), or *ISD Global's* pilot program *Counter Conversations* (Jacob et al. 2018). Purpose-built systems may

offer higher levels of control over data security, interaction architecture, and evaluation features. At the same time, they usually also have higher thresholds for user engagement, as (anonymous) registration or unfamiliar interfaces may discourage initial access. Nevertheless, their design allows for clearer ethical standards and greater pedagogical continuity, especially in long-term or high-risk cases.

Each platform logic affords different forms of relational work. On chat-based apps, counselors may rely more on brief, stabilising contact. In forums, they may scaffold collective learning. On algorithmic social media platforms, strategic messaging and resilience narratives may take center stage. Recognising these differences allows for pedagogical intentionality: choosing the right tone, tempo, and intervention form to suit each digital environment.

In addition, user groups frequently migrate between platforms, driven by generational shifts and emerging trends. Audiences once reachable on *Facebook* later moved to *Instagram* and now predominantly engage via *TikTok* or *Discord*, each platform introducing its own norms, aesthetics, and communicative logic. As a result, practitioners must continuously adapt, developing new platform literacies and adjusting their strategies in order to remain effective. This ongoing migration underscores the importance of flexibility, continual learning, and platform-specific expertise in digital counselling. Thus, platform logic should not be seen as deterministic. Creative workarounds, such as using *Instagram* stories as trigger and access points for engaging audiences in spontaneous, short-term conversations, embedding surveys into *Discord*, or linking *YouTube* videos to live chat options, demonstrate how practitioners can repurpose platform affordances for meaningful engagement. Ultimately, understanding platform logic as a co-constitutive element of digital counselling is key to realistic planning, practitioner training, and evaluation.

3.5 Counselling Diverse Audiences, Diverse Needs: Mapping Digital Counselling Users

Digital counselling in P/CVE spans a wide and heterogeneous set of user groups. Contrary to common assumptions, these are not limited to young people at risk of radicalisation.¹ Instead, digital services engage with a layered spectrum of users, each with their own distinct needs, motivations, and relational dynamics. Understanding this spectrum is crucial for tailoring communication strategies, safeguarding approaches, and evaluating impact.

1. At-Risk Individuals and Vulnerable Individuals

Digital counselling services often support to people in moments of personal instability, periods marked by emotional overload, identity struggles, or social disconnection. These 'frag-

¹ Interestingly, a recent survey conducted on online gaming shows the average age of online gamers globally to be 41, with 48 % identifying as female, 51 % as male, and 1 % as non-binary/prefer not to say (ESA 2025:7).

ile phases' do not always indicate radicalisation, but they can create openings for extremist narratives to take hold. In such cases, counsellors offer stabilising conversations, crisis support, and referral to longer-term help systems.

At the same time, counsellors engage with individuals showing early signs of radicalisation, for example through ideological references, withdrawal from social environments, or hostility toward perceived out-groups, as well as with individuals who espouse violent extremism. These cases require careful navigation between building trust and setting boundaries. Digital formats can help lower the threshold for first contact in both situations, especially when anonymity is preserved.

By clearly distinguishing between vulnerability and risk, practitioners can better tailor their interventions, offering either crisis-oriented stabilisation or more targeted distancing or deradicalisation support. In practice, these categories often overlap, and counsellors must adapt dynamically as new signals emerge.

2. Concerned Others: The Social Environment

Another core target group consists of individuals in the social environment of potentially radicalised persons. For example, parents and caregivers, siblings, teachers, mentors, or friends seek information and support in order to understand warning signs, react appropriately, or de-escalate conflicts. These users require psychoeducational and emotional support. Their sense of urgency is often high, but their understanding of digital subcultures or extremist codes may be limited. Counselling here becomes an act of translation between the lived realities of the social environment and the at-risk person they seek to help.

3. Bystanders and Helpers in Unstructured Spaces

In digital spaces such as gaming forums, Discord servers, or comment sections, peers with some awareness of potentially harmful dynamics in these spaces (sometimes dubbed 'bystanders') and other informal helpers often play a crucial role in de-escalating harmful conversations or guiding peers toward support. These individuals typically do not belong to formal counselling structures, but they can act as important bridging figures, especially when they are equipped with basic knowledge of warning signs and response strategies. To strengthen this potential, some initiatives have included bystanders and peer actors in multiplier trainings, offering them practical tools to recognise critical situations, respond constructively, and refer individuals to professional help if needed. These trainings help anchor a culture of digital co-responsibility in unstructured or semi-public online spaces.

4. Professional and Peer Support

Finally, a growing share of digital counselling is directed at professionals, teachers, social workers, NGO staff, or peer moderators. They seek short- or long-term expert advice, external resources, or simply someone with whom they can discuss difficult cases. This is not direct counselling, but rather a kind of second-degree counselling for people who are supporting others. Such interactions are crucial for sustaining preventive ecosystems; they strengthen local actors' capacity, reinforce common standards, and distribute responsibility beyond specialised teams.

Pedagogical and Evaluative Implications: The diversity of user groups requires differentiated approaches, not only in counselling methods but also with regard to evaluation. Different target groups entail different counselling objectives and modalities, resulting in entirely different counselling approaches that cannot be subsumed under a one-size-fits-all evaluation model.

Moreover, user boundaries are often fluid, with priorities and objectives evolving over time. Someone seeking brief information today may return for longer-term support weeks later, or refer someone from their social environment as a direct counselling client, while wishing to continue receiving support themselves. Others may never identify themselves as clients but can still change their behaviour based on content or conversations observed in their digital ecosystems. Evaluation models need to reflect these shifting roles, capturing impact beyond direct interaction.

In sum, digital counselling in P/CVE operates not as a linear pipeline but as a relational ecosystem. Recognising this complexity is key and directly informs pedagogical intentions and evaluation practices.

4. When Measurement Meets Reality: Rethinking Evaluation in Digital Interventions and Counselling in P/CVE

While digital counselling initiatives continue to multiply, evolve, and adapt, they are also faced with increasing demands for evidence concerning issues such as effectiveness, accountability, and quality assurance. These questions, ranging from attribution of impact to privacy-compliant monitoring, will be discussed in detail in this chapter. The institutional logic of context-shaping counselling engagements, for example whether they are NGO-led vs. state-anchored, outreach vs. reactive, peer-based vs. expert-driven, impacts what kinds of evaluation are possible, meaningful, and ethically sound. The results of this chapter are rooted in 8 expert interviews and two digital focus groups (see chapters 2.2 and 2.3). To provide practical insights that are usually hidden from administrators planning for eval-

uation, as well as a clear reasoning for the suggestions offered below, each subchapter frames a practical challenge, followed by potential evaluative responses and recommendations for tackling them.

The next chapter turns to a key question: How do we evaluate something that is constantly shifting and that also involves a degree of ambiguity in terms of how digital counselling in P/CVE can be conceptualised? This section dives into the real-life challenges that practitioners and evaluators face in these complex digital spaces and explains why flexible, adaptive evaluation methods are not just helpful, but essential in this realm.

4.1 Assessing Levels of Professionalism: Developing Stance and Standards

a) The Practical Challenge

Digital P/CVE counselling is not static. It evolves through dialogue, experimentation, and formative evaluation. Practitioners must constantly re-assess target groups, emotional dynamics, and communicative boundaries, technical challenges, as well as safety and data protection, and they require substantial flexibility.

In the early stages of digital P/CVE interventions leading to counselling, there were no clear references or professional standards to build on. While forms of classic counselling that were simply conducted via digital or other forms of remote media have a slightly longer tradition, many P/CVE approaches in this realm aimed not only to shift counselling to alternative media formats, but also to explore new ways of reaching target groups online and to convince them to participate in counselling. As a result, initial strategies often involved activist-like approaches, such as setting up fake profiles and trying to gain access to digital environments in order to counter extremist narratives. These first steps towards digital counselling in P/CVE were quite improvised:

‘We created fake profiles and tried to enter peer groups. It was anarchic and had little to do with professional counselling. It was all quite undefined, more like a social form of activism than structured counselling.’ (FGG 2, Pos. 76)

At that time, this particular P/CVE field was largely shaped by researchers with limited background in psychosocial work or counselling, requiring a complex transfer of knowledge and perspective. These beginnings lacked standards and professional expertise, often conducting immediate counter-speech in response to extremist content rather than offering actual psychosocial support. Gradually, however, a field-specific professional identity evolved:

'The first big episode was developing a stance and figuring out what standards should apply. The field was dominated by academics who didn't understand social pedagogy.' (FGG2, Pos. 77)

Only gradually did more coherent concepts emerge, described in brochures and literature, and with them the development of distinct strands such as content-driven and relationship-driven work. What is now widely understood within the field of P/CVE as a systemic, professional stance-based approach to digital counselling only became consensus over time. Therefore, the first major milestone was the development of an actual approach and professional stance, grounded in the respect for autonomy and fostering voluntary participation, by building on other forms of psychosocial counselling, especially social work and civic education. This professionalisation was informed by processes of internal reflection, practice-based learning, and formative evaluations, which helped establish criteria, language, and formats. As a result, a clear understanding of the need for defined goals, as well as professional standards and stances have become the rule, rather than the exception:

'If the internal goal isn't clear, why we're doing this and what for, then to me that's a red line.' (FGG2, Pos. 78)

'If these questions of why we are reaching out, what the goal is, and what we hope to achieve aren't clarified within the organisation, then to me that's a no-go. It opens the door to abusing power, telling people what's right or wrong, which for me is no longer counselling. That's no longer social pedagogy.' (FGG 2, Pos. 79)

This aligns with a systemic and resource-oriented understanding of social work, where trust, relationship-building, and contextual sensitivity are all central values.

b) Potential Evaluative Responses

The development of this field was marked by activist approaches that only gradually transformed into intervention and counselling formats guided by an expressed professional stance and standards. Some projects may still be in the process of transformation along these lines. To adequately understand and guide these developments, formative evaluations in general, and developmental approaches in particular, can be very useful.

4.2 Reaching Users Online: Why Evaluation Must Adapt

a) The Practical Challenge

Nowadays, practitioners describe an evolving intervention landscape where content production increasingly intersects with traditional counselling roles. Content creation has become one of the most commonly used pillars for establishing credibility and visibility,

creating a necessary access pathway between potential clients and counselling projects. This shift also impacts practitioners' professional identities, requiring them to engage with influencer-like dynamics. Questions arise about the frequency, ethics, relevance, and type of content production needed to make an impact online. Practitioners also express uncertainty regarding the extent to which they must adopt influencer characteristics in order to connect effectively with their audiences. The experts noted the practical challenges of this kind of content work (time, focus, learning, and emotional investment), while underscoring its usefulness in outreach and practice:

'Otherwise, I'm a bit of a ghost. It takes time and focus to produce meaningful content. I'm just starting to try it now to see if it helps me get noticed by potential clients.' (FGG 2, Pos. 112)

Still, practitioners also highlight the difficulty in designing content for anonymous, demographically diverse audiences:

'You're supposed to consider age when you design your content, but online, it's anonymous. You don't know who's watching or reading. That makes it really hard to target anything meaningfully.' (FGG 2, Pos. 124)

As a result, counselling strategies often rely on iterative adaptation, informed by internal monitoring and explicit or implicit forms of formative evaluation. This experience resonates with findings from the Maltese programme Kellimni (Kellimni n.d.), which has shifted from a child- and youth-centred service to a more inclusive platform accessible to people of all ages, professions, and educational backgrounds. This evolution was not only strategic but also reactive, building on their target groups' specific needs:

'We noticed that users who had accessed the service as children kept coming back as young adults. So we adapted our approach instead of excluding them.' (INT02, Pos. 6–7)

The service is now deliberately open and non-exclusive, offering anonymous and free support to anyone who needs it (INT02, Pos. 4, 6–7). Visibility is ensured through campaigns on social media channels like Facebook, Instagram, and TikTok, as well as through school visits and TV coverage. However, the actual counselling work takes place on a separate chat platform, not within the social media environments (INT02, Pos. 16–19). This responsiveness extended to thematic adjustments. For instance, earlier issues like school stress and bullying gave way to more complex themes such as abuse, suicidal ideation, and loneliness, which the project responded to through expanded services like restorative practices (INT02, Pos. 24). Monitoring plays a central role in understanding who accesses the service and due to which concerns they do so.

Furthermore, practitioners noted limited control over interactions. Digital counselling heightens this fragility: clients may disappear, sessions may end without closure, and relational cues are hard to decipher. Practitioners therefore described wishing they had more insight: 'We'd love to analyse what kinds of content we create and what kind of reactions we get' (FGG2, Pos. 133). Abrupt conversation terminations (e.g., the client leaving the chat without notice) frequently occur, posing questions about whether these are influenced by the counsellor or by external factors. There is also widespread demand for qualitative analysis of interaction data that can directly inform practice, examining communication styles, target-group reach, and triggers leading to conversation discontinuation. The anonymity and distance of digital counselling can, however, also support disclosure. For example, ExitUSA initially started out debating offline hubs but rather shifted to digital-only models due to budgetary constraints and geographic challenges. Surprisingly, they observed that clients seemed to be more open via email and phone than in-person, suggesting that anonymity lowered the threshold for sharing difficult or shame-laden experiences:

'I saw how potent it was. People opened up more than I expected, maybe because they didn't have to look someone in the eye when sharing something shameful.'
(INT01, Pos. 53–55)

b) Potential Evaluative Responses

Formative evaluation can play a pivotal role in evaluating such complex processes, including by supporting reflection on professional boundaries, authenticity, and effectiveness in digital self-presentation. It can help teams reflect on how content is perceived, whether it resonates with intended audiences, and how to adjust narratives or formats accordingly. It may also help clarify what types of content and topics are most impactful, and for whom. This would be in line with practitioners' rejection of one-size-fits-all approaches to intervention and counselling, stating instead a clear need for individualised offers and service selection (INT02, Pos. 49), which also requires more flexibility in evaluation approaches as well. Furthermore, a formative evaluation approach would pay credit to the ongoing, experimental nature of many approaches, which, in the context of fast-paced digital ecosystems, are required to constantly test new themes, topics, and ways to engage target groups based on their individual needs. For example, one project introduced theological spaces for ideologically anchored clients: 'We deal with individuals who are heavily ideologised via religious texts. You cannot not talk about religion with them' (FGG2, Pos. 134).

Rather than simply collecting quantitative metrics, such as reach, interaction rates, et cetera, the incorporation of formative elements would ensure an evaluation's relevance to practitioners, systematically supporting them when tackling complex challenges.

In summary, formative evaluations should address the complexities of digital interactions, considering practical resource limitations, the blurred boundaries between content creation and counselling, and the nuanced understanding required for fostering meaningful online engagements.

- Formative evaluations are essential due to the dynamic, unpredictable nature of digital counselling settings.
- Continuous reflection and adaptive adjustments help counsellors effectively respond to emerging needs and target groups.

4.3 Evaluating Impact: Between Expectations and Evidence

a) The Practical Challenge

Evaluating the impact of digital counselling services remains one of the most contested and complex dimensions of this rapidly evolving field. As outlined earlier, many projects operate with limited resources, making it difficult to implement comprehensive *Monitoring, Evaluation and Learning* (MEL) structures. The perceived purpose of evaluation also plays a role. For many, it may be less about learning and more about legitimising project funding. Practitioners worry that this instrumental logic pushes them to chase metrics that are only loosely, if at all, connected to meaningful change. This particular challenge holds true for both offline and digital contexts, albeit with an added layer of complexity in the case of the latter.

In digital contexts, already complex challenges (e.g., the varying expectations of funders, practitioners, and evaluators) are further entangled with issues such as conceptual ambiguities, technical conditions, and platform-specific affordances. Among all these actors, the desire to evaluate the effects of digital counselling goes far beyond merely counting sessions or measuring audience engagement metrics. That said, across all 14 projects within which interviews or focus groups were conducted for this study, not a single one had been subjected to a formal impact evaluation. The reasons for this are multifaceted. Structural constraints, such as limited time, staff, or funding, are one issue. But deeper questions, tied to overarching debates around measuring the impact of social interventions, also remain unresolved and unfold in the digital section: What does 'impact' mean in (digital) counselling? How can it be detected? Who defines it? Practitioners report that the effects of their work often become visible in subtle and unintended ways:

'Sometimes a passing comment makes all the difference for a client, and we as practitioners don't even intend it. So how can we measure that? And do we really expect clients to come back after a year to tell us what changed in their lives? Did it change because of us, or because of something else? These are not easy questions.' (FGG1, Pos. 58)

In non-digital settings, however rare, evaluations can include follow-up interviews or longer-term observation, provided sufficient funding and appropriate time-frames are made available. In digital spaces, however, high levels of anonymity and the one-off nature of some interactions further impose limitations on such approaches. This is especially true for interventions on open-access platforms or low-threshold counselling services, such as anonymised chats, among others.

This challenge is exacerbated by the nature of platform interactions. The quality, tone, and dynamic of exchanges differ significantly between platforms such as *TikTok*, *Instagram*, *Discord* or *Twitch*. For instance, *Discord* allows for slower-paced, structured discussions with moderation tools that limit spam and promote reflective exchange:

‘You can make it so that each person can only post once every five minutes... the discussion becomes more focused, and that’s incredibly effective.’ (FGG2, Pos. 74)

The textual depth and relational patterns fostered in such an environment are vastly different from the quick-response, attention-fragmenting mechanics of algorithm-driven content platforms. Digital counselling projects often operate across this spectrum of platforms. On the one end, there are direct counselling settings, such as one-on-one chats, where trust-building and emotional reflection are central. On the other end, there are campaign-like intervention formats that disseminate content into broader public spheres, attempting to challenge filter bubbles or introduce alternative narratives, in order to create interest in and access paths to counselling settings.

The diversity of goals and intertwined mechanisms also demands a pluralism of indicators. Funders often seek numeric output: reach, clicks, participant numbers. Counselling practitioners, by contrast, focus on relational depth and qualitative transformation: Is a client more reflective, more stable, more socially embedded? Researchers, meanwhile, often look for methodologically sound comparisons and seek to approximate causality or correlation across cases.

This divergence in evaluative logics and professional or institutional priorities is rarely resolved. Projects often find themselves stretched between producing strategic reports to satisfy funders while engaging in reflective self-evaluation for internal learning:

‘It’s mainly a numbers game. We can show that our project was well received, but it says little about the content quality of our workshops.’ (FGG1, Pos. 63/65)

While some practitioners reported positive experiences with funders open to qualitative insights, others noted that rigid output requirements can compromise actual quality:

'One year, our report highlighted that we were invited onto a popular gamer stream. That looked great in numbers. But what does that say about the quality of our counselling work? Nothing, really.' (FGG1, Pos. 65)

'If you're under pressure to produce content every week... the quality of the counselling suffers.' (FGG2, Pos. 94)

A strong pressure to deliver quantifiable success stories remains, sometimes at the cost of professional quality and leading to the distortion of work priorities. While metrics such as click rates or the number of counselling sessions can paint a picture of reach, its significance to assess resonance and relevance regarding attitudinal and/or behavioural change is limited at best.

Still, not everything is bleak. Some practitioners reported that, over time, some funders began to welcome qualitative case reports and open formats for storytelling, especially once the limits of standardised reporting became evident. One participant explained:

'At some point, even our funders asked us to just describe what happens behind the scenes. What kind of conversations are we having? What works? What doesn't? That was a shift, a welcomed one.' (FGG1, Pos. 67)

Such shifts also reflect the growing experience and professionalisation within the field. As organisations become more confident in their own methods, they are increasingly able to make a case for nuanced reporting, balancing numeric data with qualitative insights.

Another major challenge lies in the tracking of longitudinal effects. Anonymous or short-term digital counselling services face particularly difficult obstacles in this regard: Clients may not be reachable for follow-up, data protection rules may restrict data linkage, and even when follow-up is attempted, it often remains unclear how changes in a client's life can be attributed to the intervention alone – a classic problem that is well-known to social intervention practitioners both off- and online:

'Clients aren't always willing to participate in a follow-up a year later. And even if they do – how do we know it was us, and not some coincidental positive life development?' (FGG1, Pos. 58)

b) Potential Evaluative Responses

Nevertheless, it would be a mistake to conclude that evaluation is simply not possible. One way to start overcoming these challenges would be a shift in epistemology: from definitive proof to plausible interpretation, from generalisable findings to situated understandings. Some projects have started to explore more adaptive methods. These include qualitative

case tracking, relational diagnostics, and the use of platform-specific engagement metrics that go beyond reach, such as the content of replies, patterns of sustained engagement, and shifts in client language use over time. One practitioner described observing how ‘certain conversations change over time... even small shifts in attitudes, we try to pick up in qualitative analysis’ (FGG2, Pos. 80). Some teams observe whether clients’ messages change in tone and complexity throughout a counselling relationship, using this as a heuristic for cognitive and emotional development. While such methods are not without their challenges, they can open paths towards understanding dynamics that are invisible to conventional metrics.

The tension between the demand for accountability and the need for learning runs through all evaluation efforts. While practitioners acknowledge the legitimacy of funders’ needs to justify spending, they stress that impact evaluation must not become a tick-box exercise:

‘Evaluation is legitimate – of course public funding requires justification – but for development purposes, these metrics are only somewhat useful’ (FGG1, Pos. 60). I understand why public funders want evidence. But if we focus only on outcomes that are easy to quantify, we miss the actual work that’s being done. Qualitative stories matter, too, especially for improving practice.’ (FGG1, Pos. 60)

To bridge these different priorities, shared learning environments should be established, where both funders and service providers exchange insights outside of formal reporting structures. Indeed, the most promising developments point towards mixed-method approaches which consider both numbers and narratives, structure and surprise:

‘We tried to show that something had changed qualitatively – even if the numbers were small. But you have to make that appealing to funders too, even if it’s only one or two people. When we can show that someone got redirected to further help, that counts. But we also need better ways to visualise these effects.’ (FGG2, Pos. 18)

In the end, quantitative and qualitative indicators must be understood as complementary rather than contradictory. While numbers can reveal trends, they rarely illuminate context. The success of a counselling interaction, particularly online, often hinges on trust, tone, and timing. As such, the quality of the relationship must remain a central dimension in any impact assessment.

‘I’d love to see more qualitative evaluations, like we have in social work. Not just numbers, but what really happened. What were the needs? Were they met? Patterns can emerge, but only if we look deeply enough.’ (FGG2, Pos. 145)

Ultimately, digital counselling services operate in an environment marked by high volatility, emotional intensity, and limited continuity. To grasp their impact, we need tools that are

equally sensitive, flexible, and varied. Evaluation in this field must reflect the ambiguity and complexity of human development processes mediated through digital infrastructures. As the field continues to evolve, one thing becomes clear: it is not only about measuring impact, but also about cultivating an evaluative culture that respects the depth and diversity of digital relational work. Whether these insights will translate into more sustainable evaluation practices remains to be seen, but the findings suggest a growing awareness that, if digital counselling is to be taken seriously, so must the complexity of its impacts.

- Measuring impact in digital counselling is complex due to anonymity, short interactions, and diverse platform dynamics.
- A mixed-methods approach combining qualitative and quantitative data is necessary to capture meaningful changes and avoid oversimplification.

4.4 Bound by the Platform: How Tech Shapes Impact

a) The Practical Challenge

Digital counselling operates within the structural boundaries and affordances of the platforms it inhabits. These platform dynamics, ranging from the technical architecture to the willingness of the tech industry to engage with external P/CVE or social service providers, profoundly shape what types of interaction are possible, how relationships are formed, and ultimately, what kinds of outcomes can be evaluated.

The bifurcation of digital counselling into outreach-based and passive models also plays a central role in this structure. Outreach projects rely on platforms as direct engagement spaces, sometimes as the primary site of contact. Passive models, by contrast, often use external platforms only as transitional zones, with the goal of moving clients into secured, project-owned systems. This decision is not trivial: it alters the conditions under which impact can be measured, confidentiality ensured, and relationships cultivated. In practice, projects often develop hybrid models, combining asynchronous tools like secure messaging applications with live interactions on external platforms.

Even when real-time conversations are possible, they are shaped by underlying platform architectures. For instance, *Discord's* previously mentioned moderation tools, limiting users to posting once every five minutes, can fundamentally alter the quality of interaction. These kinds of settings were described as effective for reducing spam and fostering more meaningful discussion, thereby creating a communicative climate that is more conducive to trust and self-reflection (FGG2, Pos. 74).

Practitioners also observe that platforms shape the type of conversation that becomes possible:

‘What really makes the difference is the interaction pattern. On Discord, the quality of conversation is entirely different than on TikTok. We see complex, sustained argumentation, more like a discussion, whereas on TikTok it’s more like rhetorical fencing. You get short, reactive bursts, not dialogue.’ (FGG2, Pos. 74)

This difference has profound implications. Platforms not only serve as infrastructure; they actively co-produce communication formats, influencing whether relationships can emerge at all, and in what form. While online chats may support comprehensive reflective processes, interventions in comment sections on fast-paced platforms rarely do, but rather provide opportunities for creating initial interest in support structures.

However, not all platforms are equally cooperative. Practitioners report that while some services offer structured exchange and active communication with platform moderators, others present severe obstacles:

‘Steam was a huge challenge. The guidelines explicitly state that external moderation or so-called ‘wannabe counselling’ is unwelcome. We felt that immediately. We couldn’t join the discussion meaningfully; in fact, we were often blocked or restricted. Digital streetwork was almost impossible there, so we moved to Twitch and Discord instead.’ (FGG1, Pos. 38)

In contrast, *Discord* provided openings: One project reported that the moderation staff had been open to their suggestions. When profiles with clearly criminal or extremist content were flagged, the team responded positively. This kind of exchange was described as particularly valuable (FGG1, Pos. 38). Such experiences showcase how space and policy can shape counselling services both in online and offline spaces.

In some cases, practitioners suspect that their content and messages were rendered invisible without notification, a phenomenon they refer to as ‘shadow banning’. The suspicion arose, for instance, when posts received no reactions at all despite having worked well previously. Since there was no warning or flag, practitioners could only guess whether a shadow ban had occurred (FGG2, Pos. 33). This lack of transparency creates a need for practitioners to explore and invest serious resources into exploring the kind of messaging that would be allowed by certain platforms:

‘We avoided flagged terms, tried different formats, adjusted tone, but still, posts would vanish. There’s no explanation, no rationale. You’re just left guessing.’ (FGG2, Pos. 35)

Even after direct conversations with platform representatives, clarity often remains elusive:

'TikTok said we were doing something wrong, but they wouldn't say what. Things did improve after we talked, but we still have no idea what changed.' (FGG2, Pos. 37)

Such unpredictability forces practitioners into constant workaround mode. As one participant described it: 'Sometimes, the only option is to scrap the profile and start over. But there's no guarantee the new one will work either. It's a gamble' (FGG2, Pos. 47). These technical instabilities undermine consistency and evaluation alike, making it difficult to attribute success or failure to any single intervention. Even seemingly minor bugs carry weight:

'One of my colleagues wasn't notified of replies to her comments. She had to manually search for them. That tripled her workload. And it happens mostly on TikTok.' (FGG2, Pos. 47)

These invisible frictions accumulate, distorting the perceived responsiveness of counseling offers and eroding practitioner motivation. Practitioners have also observed a gradual shift in how some platforms engage:

'It was fantastic that we even managed to get in touch with someone at TikTok. Things did improve afterwards, but we still don't know why. Was it something we changed? Or did they alter their systems? We still don't really know.' (FGG2, Pos. 43)

This illustrates a lack of clarity even when progress is made, and also how difficult it can be to translate informal improvements into stable practice. At the same time, the lack of transparency from platforms adds another layer of difficulty. Especially on platforms like *TikTok* and *Instagram*, practitioners report that they have no control over which messages get delivered. As one expert recounted, *TikTok* responded vaguely to a blocked campaign, saying only that something had been done wrong, without offering clear guidance for improvement (FGG2, Pos. 47).

International comparison revealed different models. One expert with experience working in the U.S. clearly described that they had no illusions about the nature of the platform they used:

'It's simple: they're a company, and we pay to use their service. There's no other cooperation, no sponsorship, no dialogue.' (INT01, Pos. 36)

This illustrates a more transactional approach, one in which counselling merely operates within platform constraints, without any further structural collaboration.

In some cases, projects sidestepped uncooperative and intransparent platforms by building informal networks. One practitioner explained that when their project launched, they initially focused heavily on field networking, directly approaching agencies and individual influencers. This strategy led to invitations to participate in live streams and, gradually, to an expanding network of collaborators (FGG1, Pos. 20). These relational strategies can create new access points outside of platform infrastructures, and can also reflect a broader shift toward peer- and community-driven outreach.

Ultimately, the diversity of digital platforms and their embedded logics present both a challenge and an opportunity. On the one hand, practitioners must continually adapt to shifting affordances, opaque content rules, and variable levels of cooperation. On the other, these platforms also allow for creative infrastructural interventions and hybrid solutions that shape not only the counselling interaction, but also how success and change are understood and evaluated. As one practitioner concluded: 'Platforms are not just spaces we use, they're transmission media. They shape what's *possible*, and we have to understand them that way' (FGG1, Pos. 36).

b) Potential Evaluative Responses

Evaluations conducted in this space have to take these technological and conceptual platform boundaries seriously, by carefully assessing the data collection possibilities and limitations while ensuring that the analysis takes into account the intense efforts necessary to conduct professional and ethical counselling in these spaces. The constant (re-)development of practical approaches necessary to continue functioning in these settings is not just a side note, but an incremental part of such practical approaches and needs to be treated as such. Formats such as process and developmental evaluations hold significant potential to support practitioners in their work and contribute to more effective implementation of practical approaches.

- Both evaluations and interventions are significantly shaped by platform-specific constraints, such as algorithmic limitations and moderation policies.
- Platform-dependent methods are required, emphasising adaptability and collaboration with platforms.
- Process and developmental evaluation approaches are particularly suitable for supporting practitioners and contributing to evidence-based practice.

4.5 Adapting on the Fly: Digital Counselling Without a Roadmap

a) The Practical Challenge

Digital P/CVE counselling practices are marked not only by planned, structured interventions, but also by a high level of spontaneous emergence. New needs, topics, platforms, audiences, and many more challenges often require practitioners to adapt quickly, creating new activities that were not foreseen in the original project design. One practitioner summed it up: 'It's always reactive. We didn't even know what we needed until it happened and by then, we had already moved on' (FGG2, Pos. 33).

Projects regularly find themselves improvising around shifting conditions. These might be platform-related, such as unexpected algorithmic changes, content removals, and other issues discussed in the previous section, or even the need to abandon a space altogether (for example, due to hostile user reactions or platform policy changes). In some cases, entire outreach workflows are shaped by these contingencies. One practitioner reflected that they spend most of their time navigating *TikTok*, allowing the algorithm to guide them to content worth engaging with. Once they find promising conversations, they interact and see if dialogue unfolds (FGG2, Pos. 68–70), leading to high levels of volatility in daily workflows, depending on the specific topics, audiences, challenges encountered:

'We used to post content and wait. Now we follow the algorithm, comment, interact, and hope to stay visible. That's our daily workflow.' (FGG2, Pos. 68–70)

Another practitioner described how hard it is to maintain visibility: They had to start creating more regular and strategic content just to be noticed. Before that, they felt they had rarely been perceived by potential clients beyond occasional random posts. However, producing meaningful content requires time and focus, which is a rare resource in everyday practice, especially when content production is primarily used as a means to engage audiences in dialogue, rather than being the key aim itself (FGG2, Pos. 15).

One project lead explained how the team initially tried to accompany users across various platforms, until their content suddenly became invisible on *TikTok* and interactions ceased entirely. They later learned from discussions with the platform that their account had likely been shadow banned (FGG2, Pos. 33, 35). This shows how visibility itself has become a metric of success in this space:

'We didn't even know it happened, we just thought our audience was gone.'

'If you're not seen, you can't help. So, we started posting more, testing content types, even just to stay relevant.' (FGG2, Pos. 15)

In general, the quickly emerging responses and adaptations conducted by projects are rarely random. They are often necessary to maintain contact with at-risk individuals or to preserve credibility within fast-changing subcultures. Over time, initially ad-hoc activities, such as the active monitoring of toxic debates, or the interjection of counter- and alternative narratives have grown into a core strategy without being part of the original mandate or theory of change. For example, one team described how their internal monitoring flagged hostile reactions toward black characters in gaming communities. In response, they joined comment threads and livestream chats to offer alternative views, an approach that had become routine but had not been planned originally (FGG1, Pos. 52). In many cases, the project focus significantly evolved over time. One team originally planned on sparking discussions via content production but later found that their core work had shifted toward moderating debates, promoting ambiguity tolerance, and reframing narratives within comment sections or live chats, especially through a 'one-to-many' logic (FGG2, Pos. 50).

b) Potential Evaluative Responses

In such highly volatile contexts, where predefined indicators might become irrelevant due to changing circumstances, some projects have turned to ad-hoc internal evaluations. For instance, one team used a meta-review format: they evaluated internal feedback at the end of the year to identify gaps, training needs, and improvement areas (FGG1, Pos. 54). Others emphasised self-evaluation as part of everyday practice, often conducted in parallel to formal external evaluations (FGG1, Pos. 50). Another practitioner mentioned structured annual processes:

'We had regular intervision groups, where five professionals reviewed one or two cases over 90 minutes. The discussions were then summarised and used for training design.' (FGG1, Pos. 54)

When projects are forced to reconfigure what counts as success, evaluation indicators need to be adjusted accordingly. In classical counselling logic, indicators might include stabilisation in daily life, exit from extremist networks, or strengthened pro-social ties. But in digital outreach counselling, for example, according to many of our interview partners, success meant simply remaining visible at all, maintaining a presence in hostile environments, or initiating reflection.

This high level of volatility points to a pressing need for more adaptive evaluation frameworks. Firstly, evaluation designs need to anticipate emergence. They should explicitly include mechanisms for identifying and capturing unplanned outputs and outcomes. Secondly, process-oriented evaluation should be standard, not optional. It is only by tracking how and why projects adapt that we can understand what makes them effective. Finally,

funders and evaluators alike must be prepared to value the invisible: work that keeps digital counselling present and legitimate in fast-shifting, often adversarial environments.

To support adaptive evaluation, the following indicators may help capture meaningful effects across digital P/CVE contexts. These indicators may not replace traditional metrics, but they more accurately reflect the lived complexity of digital P/CVE work:

- **Platform presence and visibility:** Frequency and consistency of engagement across key platforms; documentation of shadow banning or content suppression events.
- **Narrative interventions:** Instances where alternative or counter-narratives were placed in toxic debates; perceived tone shifts in threads.
- **User resonance:** Direct feedback from users (e.g., surveys, spontaneous comments) indicating perceived relevance or emotional impact.
- **Process adaptations:** Documentation of strategic or methodological shifts in response to platform dynamics or user behaviour.
- **Relational indicators:** Continuity of contact, user return rates, or deepening engagement over time.
- **Self-evaluation cycles:** Existence and regularity of reflective practices such as intervention or internal review sessions.
- **Qualitative traces of impact:** Practitioner observations of change, user quotes, or reconstructed interaction narratives.

In summary, digital counselling formats often face unexpected challenges which require practitioners to adapt quickly. These emergent adjustments, usually without predefined indicators, call for a flexible, process-oriented evaluation approach that can capture and assess spontaneous activities and shifts in practice.

4.6 Too Much, but Never Enough: Data Limits in Digital Counselling

a) The Practical Challenge

Digital counselling in P/CVE presents a paradox: while digital environments offer seemingly rich data streams and new evaluative possibilities, practitioners regularly face significant barriers to systematically leveraging them. These challenges range from data access limitations and technical platform constraints to ethical dilemmas around anonymity and informed consent. Practitioners are aware of their lack of influence over the handling of user data when communication takes place via public social media and messenger platforms (FGG2, Pos. 32), which poses a challenge for ethical counselling work in these spaces. Nevertheless, platforms differ significantly in their technical and ethical affordances. While *YouTube*, for example, allows data scraping and comment harvesting, Instagram

disables message archiving, while *Discord* restricts automated analytics. Interestingly, in terms of video calls, some platforms reportedly did not allow screen recording or screenshots, which professionals considered crucial for client safety and confidentiality. They emphasised their belief that these features were deliberately disabled to prevent chat content from being leaked (INT01, Pos. 43):

‘With regards to collection of personal information, because the service is anonymised, it is a bit challenging because really and truly we wouldn’t know the personal information of the client unless we need to report, and the client is giving us their personal information to report.’ (FGG1, Pos. 27)

b) Potential Evaluative Responses

While many projects reportedly set up post-counselling surveys, some noted the limitations of this approach regarding scientific robustness (FGG2, Pos. 84).

Despite constraints, some teams have developed more creative solutions for internal documentation. Some practitioners adapted existing counselling and case management tools to document basic information, communication topics, and platform-specific data. They explained that they recorded how conversations went, what was discussed, and which overarching categories applied, such as work, social issues, or health, especially on platforms where many discussions emerged organically (FGG2, Pos. 78). One team explained that they worked with software to track age, conversation content, and conversational flow. They said,

‘We also sometimes analyse screenshots of public interactions, either ones that were sent to us or ones we collected ourselves, to gain qualitative insights.’ (FGG2, Pos. 80)

To uphold data security, documentation was designed to be handled individually by each practitioner. Not all team members had access to shared data, and this decentralisation was presented as a privacy safeguard (FGG2, Pos. 80).

However, even in cases where sufficient data could be collected, many actors struggled to analyse it meaningfully. The tension between data abundance and resource scarcity was a recurring theme. While this tension is not exclusive to the online space, it is heightened in such settings:

‘No chance, that’s just too much. There are too many debates, too many streamers we could respond to (...).’ (FGG1, Pos. 56)

In general, there was also strong interest in formative, qualitative evaluation (see also chapter 4.2). One team said they wished they had more time for in-depth analysis of chat interactions and content reactions. They wanted to understand what kind of content resonated and whether their communication style was actually effective (FGG2, Pos. 98):

'We have a time problem, we're sixteen contributors across many platforms, and we just can't do in-depth analysis of interactions the way we'd like. We can't continuously gather and analyse the data we collect. That would require circulating evaluation like in a grounded theory approach. But it's not possible because we're also writing monitoring reports, briefings, and doing other content work at the same time.' (FGG2, Pos. 98)

Some teams expressed interest in comparative analysis across modalities. They said they would like to compare the outcomes of chat- versus video-based counselling formats to optimise their methods (FGG2, Pos. 46).

Beyond technical and logistical issues, ethical challenges remain central. Anonymity is often non-negotiable for target groups, but it limits the ability to maintain qualitatively deeper progress tracking:

'After three years of anonymous counselling, we realised it's really hard to write a meaningful, varied report. That's just the nature of anonymity, there's not much relevant personal data to work with.' (FGG1, Pos. 67)

Many actors had already responded by developing custom privacy plans. One team said they informed users about platform-level data practices through a notification banner, comparable to cookie notices, at the beginning of interactions on *Discord* (FGG2, Pos. 32). Others reported building their own tools and platforms to meet legal requirements. Another team explained that they could not sign third-party data processing agreements with U.S.-based service providers due to EU data protection laws, which led them to develop their own systems (FGG1, Pos. 37).

Still, platform dependency remained a structural constraint. One team explained,

'We just don't do counselling through some platforms at all. For example, on Instagram, we're reactive only, we don't initiate anything there. If someone reaches out, we try to move the conversation to Discord, where we can ensure better privacy.' (FGG2, Pos. 32)

Finally, practitioners saw ethical issues as more than a compliance requirement. Ethical evaluation was presented as a continuous negotiation between impact and protection, re-

quiring not only measurement tools but also strategic decisions around engagement, trust, and safety.

All this demonstrates that evaluation in digital P/CVE contexts must account for both technical and normative tensions. Data protection and user anonymity make it difficult to build robust evaluation datasets, for all the data allure of working on digital platforms. Future-oriented evaluation should thus combine ethical sensitivity, contextual awareness, and methodological adaptability, rather than seeking one-size-fits-all metrics for success. For this reason, clear ethical guidelines, strong data protection policies, and the development of privacy-compliant digital tools are essential.

- Technical constraints and ethical considerations strongly limit systematic data collection in digital counselling.
- Developing clear ethical guidelines and privacy-compliant tools is necessary for improving evaluative data quality.

4.7 Doing More with Less: Practical Approaches to Evaluation

a) The Practical Challenge

A central cross-cutting finding from the focus groups and interviews is the significant lack of resources for MEL measures in digital counselling projects. In practice, even simple self-evaluation measures, such as the structured documentation of contacts or interventions, are rarely implemented systematically. The reasons for this are manifold: limited time, staff shortages, a lack of training in evaluation methods, and the generally low institutional priority of MEL in the day-to-day routines of digital counselling settings.

'Honestly, I don't know how one could integrate quality measurement more into everyday work. That would be really hard to accommodate in our daily routines.'
(FGG1, Pos. 17)

Another participant pointed out that existing evaluations often do not benefit the projects themselves, but primarily serve accountability purposes instead. This leaves a key potential untapped:

'I don't think that's the only thing it [evaluation] can be, because for further development, in my view, it's only of limited help.'
(FGG1, Pos. 16)

b) Potential Evaluative Responses

The factors above result in a clear demand: External evaluations, when conducted, must be use-oriented. Their primary purpose should not be accountability to funders, but rather

the further development and the quality assurance with regard to practice. This requires close integration between evaluation teams and project leads, as well as iterative formats whose insights can be directly applied to practical adjustments. This demand aligns with the principles of formative evaluations as emphasised in both the literature and, increasingly, by practitioners.

The call for 'cooperation instead of everyone cooking their own soup' (FGG2, Pos. 102) points to a systemic challenge: Digital counselling still often operates on the margins, frequently without structural integration into existing quality development networks or interdisciplinary discourses, such as those in social work, IT, media pedagogy, or research. Building such structures would be essential not only for knowledge transfer and peer support, but also for the sustainable use and safeguarding of evaluation findings. This view is echoed in international expert interviews, where a similar lack of cross-sector collaboration was observed:

'Instead of bringing in people who really understand data, say, from the social media context, we all just cook our own soup.' (INT01, Pos. 4)

Another challenge is related to staffing: 'So much knowledge is lost due to high staff turnover, we need similar networks as in other areas of prevention' (FGG2, Pos. 103).

When projects operate without spaces for exchange and reflection, knowledge becomes fragmented, evaluation results dissipate, or even go undocumented. A collaborative, reflective approach to practice development, supported by ongoing formative evaluation, could therefore represent a promising path toward the professionalisation of digital counselling in the P/CVE field, provided adequate resources are available.

At the same time, it appears that a shift in general perception is necessary: Evaluation should no longer be understood as an additional burden, but rather as part of professional practice, and it must be designed accordingly. This requires simple, low-threshold, and time-efficient methods that can be integrated into daily routines. As one participant put it:

'I'm convinced that numbers can reveal incredibly important insights. But even the question 'What do I want to know?' isn't precise enough yet.' (FGG2, Pos. 19)

Only then can MEL, even set against limited resources, truly unfold its impact. Addressing this systematically requires practical solutions such as simplified monitoring tools, cross-institutional networking or collaborative data-sharing.

- Limited time, staffing, and MEL expertise hinder systematic evaluation in practice.

- Evaluations must become use-oriented, practically feasible, and embedded into daily routines rather than seen as additional burdens.

4.8 Lost in Translation: Funders, Researchers, and Practitioners

In the context of digital counselling and prevention work, the relationship between practitioners, funders, and evaluators is shaped by a persistent challenge: the absence of a shared vocabulary for formulating goals, describing outcomes, and constructing plausible impact models. Digital projects seem prone to a mismatch of expectations. Funders often rely on traditional evaluation logics that assume linearity and measurability, while practitioners operate in dynamic, anonymised, and often experimental settings. Practitioners frequently reported that the indicators they are expected to deliver offer little meaningful insight into the actual impact of their work:

‘It’s all numbers, how many people reached, how many sessions held. I understand why funders want this, but it tells only a small part of the story.’ (FGG1, Pos. 18)

Many projects in this realm appear to begin without a detailed plan for how to tie their theory of change to robust indicators. Instead, they evolve iteratively, guided by needs and interaction patterns in digital spaces. As one interviewee put it:

‘We started out with numbers and reported those. But after three years, we realised anonymous counselling generates very little evaluable data. You don’t even know if you’re really making a difference, at least not in the ways that show up in spreadsheets.’ (FGG1, Pos. 18)

What emerges from these accounts is not simply a methodological challenge, but a deeper communication gap. While funders demand accountability in the form of metrics, practitioners are seeking ways to narrate the complex, adaptive work they are doing. Evaluation can play a bridging role here, if it is reimagined as a dialogical and learning-oriented process:

‘Honestly, the best evaluations were always the ones where we had regular check-ins with the funder, outside of reporting cycles. That’s when you can explain what’s really happening and why some indicators are missing or irrelevant.’ (FGG1, Pos. 18)

Encouragingly, some funders have shown openness to such insights. One team recalled that, after years of struggling with generic reporting templates, their advisory board unexpectedly asked for more reflective, qualitative insights:

'They told us, 'Don't just give us numbers. Tell us what's going on behind the scenes, how you work, how you make decisions.' That shift meant a lot to us.' (FGG1, Pos. 18)

Other projects reported similar challenges: One initiative found that due to anonymised chat formats, they had little relevant data to report after several years, despite meaningful work being done. Instead, funders eventually asked them to describe internal dynamics and emerging trends in qualitative terms, highlighting shifts in youth behaviour or parenting strategies rather than relying solely on statistics (FGG1, Pos. 67). One participant reflected: 'I'd see it as a project on its own, to develop methods and learn what really works. It's not just reporting, it's research' (FGG1, Pos. 45).

Practitioners also note that more sophisticated data analyses would be desirable: to examine comment structures, content engagement, and how different counselling personalities shape communication outcomes. However, time constraints and limited staffing often reduce such ambitions to wishful thinking: 'We definitely prefer qualitative evaluation. But we were overwhelmed. We needed external support for that' (INT01, Pos. 85). Others point to a persistent tension between funder expectations and professional standards. As one expert explained, statistics were required for reporting purposes, but quality of care was the actual priority. 'For me, the numbers are secondary. What matters is that our counsellors are well-trained and stay up to date with best practices' (FGG1, Pos. 51). This emphasis on quality is also linked to the highly individualised nature of the work and the need to prioritise process-orientation. Long-term evaluation partnerships offer one possible solution:

'The difference is huge. If someone follows the work long enough, you understand not just the results, but how they come about' (INT01, Pos. 93).

In some instances, evaluation frameworks were also co-developed with universities and civil society partners, allowing for robust methods while also fostering mutual learning across analogue and digital contexts (FGG1, Pos. 50, 54). Such co-developments and other types of mid- to long-term evaluation partnerships between practitioners, funders, and evaluators well-versed in the subject matter offer one potential solution, but require time.

To achieve this, a shared language is needed, accommodating the constraints of anonymity, the unpredictability of digital outreach, and the adaptive logic of digital engagement. This language does not yet fully exist. It can, however, be argued that building it remains one of the most important tasks for future evaluation frameworks in the digital realm.

- Differing expectations and terminologies between practitioners, evaluators, and funders create barriers to effective evaluation.

- Regular, structured dialogue and the development of a shared evaluation language are essential for aligning stakeholder expectations.

5. Looking towards the Future: From Hype to Help? The Potential of AI in Digital Counselling

Artificial intelligence (AI) remains largely underused in digital P/CVE work. That said, its potential for both counselling and evaluation is considerable and should not be overlooked.

In the short term, one practical step could be to test AI tools that help spot high-risk conversations early, while making sure data is handled responsibly and ethical standards are strictly followed. Looking a bit further ahead, teaming up with institutions that have experience in AI could help in developing tailored evaluation tools that really fit the needs of digital counselling work. However, these opportunities also raise serious concerns about feasibility, ethics, and data security.

Applied Potential: From Chatbots to Real-Time Analysis. AI may assist in identifying needs and risks early on, for example through automated analysis of text, speech, or engagement patterns. These tools could flag content shifts, identify recurring themes, or highlight urgent cases. For instance, work led by *Moonshot* leverages AI to detect at-risk users and match them with tailored prevention offers, using real-time behavioural data from search engines and platforms (Moonshot 2025).

AI can also structure case data and enable low-threshold entry points via chatbots or FAQ-driven triage. In this sense, AI may 'buy time' for practitioners: administrative or preliminary work is partly automated, enabling staff to focus on substantive counselling. Some practitioners welcome these possibilities, especially where documentation remains a resource bottleneck:

'We're stuck between doing the actual counselling and documenting it. If something could help structure that, automatically, even better, that would be huge.'
(FGG2, Pos. 98)

While AI could be developed into a useful tool to decrease practitioner team burdens, another much-debated use of AI remains practically and ethically problematic: the use of chatbots.

Complex emotional dynamics, situational nuance, and trust-building cannot yet be replicated by current AI systems: A chatbot might be able to ask questions, but it cannot hear hesitation or a sudden, slight shift in tone while answering. To date, they also tend to repeatedly affirm users, even in cases of potentially harmful or misguided statements (Abrams 2025). In the

context of psychotherapy in the U.S., the *American Psychological Association* has called for firm safeguards to avoid future harms, following serious violent incidents related to chatbots that pretended to be licensed therapists (Abrams 2025). Simultaneously, they acknowledged that professionally tested bots grounded in psychological research might eventually be able to contribute to the field and fill in some gaps (Abrams 2025). For now, however, AI should assist, not replace, the human connection at the heart of effective counselling.

Evaluation: AI as Analytical Support, Not Replacement. In evaluation, AI could help process large volumes of qualitative and quantitative data, such as chat logs or platform engagement. One expert described the situation as follows:

‘We’d love to do more with our chat data, compare styles, outcomes, reactions. But honestly, we can’t keep up. Not without help.’ (FGG2, Pos. 46)

Here, AI could help extract and visualise key trends across multiple services or formats, enhancing comparability and learning. Some also see potential in AI-assisted theory of change development, flagging logical gaps in project design or mismatches between inputs and intended effects. Others remain cautious:

‘If you don’t know what you’re looking for, data won’t save you. It just adds noise.’ (INT02, Pos. 38)

Governance and Ethics: Striking the Balance. Notwithstanding these advantages, AI also introduces new layers of complexity. Data protection is not just a legal requirement; it is a cornerstone of trust in digital counselling. One practitioner noted:

‘We want to protect the client. That’s the first rule. But without data, how can I prove that what I do works?’ (INT02, Pos. 56)

Any use of AI must therefore comply with strict privacy standards. Open-source solutions or collaborations with vetted tech partners could reduce costs and promote transparency. Some practitioners proposed public-private partnerships for accessing secure analytics without compromising user integrity. Furthermore, the risk of algorithmic bias, false positives, or unintended exclusion is particularly problematic in a context that demands fairness, inclusion, and contextual sensitivity.

Toward Responsible AI Integration. For AI to become a meaningful asset, several conditions must be met:

- Transparent and participatory development processes

- Ethical guidelines co-created with practitioners and their respective professional college or professional associations, if applicable
- Ongoing training and risk monitoring
- Complementary (not substitutive) use in counselling
- Data minimisation and secure processing infrastructure

Ultimately, as one respondent summarised: ‘AI could help, but it must serve the counselling logic, not replace it’ (FGG2, Pos. 99). By approaching AI not as a shortcut but as a supportive tool for structured documentation, pattern recognition, and theory refinement, P/CVE actors can indeed unlock new possibilities, while remaining accountable to the human core of their work.

6. What We’ve Learned – And Where to Go Next

Digital counselling has become a vital and much-discussed part of prevention work. But the very things that make it so powerful – its speed, potential anonymity, and broad reach – also make it hard to evaluate in traditional ways. Often, its impact happens quietly and invisibly: in fleeting chats, on different platforms, behind usernames that can change. Counsellors often witness moments of change that leave no digital trace, and yet they can be precisely the impacts that matter most.

Our analysis highlights four central insights:

1. **Plural evaluation purposes.** Process reviews, outcome tracking, and learning-oriented evaluations each serve a different purpose. In digital spaces, flexible, formative approaches work best; they protect user anonymity, adapt quickly, and are able to capture process quality. Emerging user generations and their migration between platforms also present an ongoing evaluative challenge. Each new generation brings with it distinct interaction styles, requiring counsellors and evaluators to continuously adapt to rapidly evolving platform dynamics.
2. **Structural asymmetries.** Digital counselling services are heavily shaped and impacted by the technical and structural features of the platforms they rely on. Challenges like algorithmic invisibility (e.g., shadow banning) or strict moderation rules make long-term, systematic evaluation difficult. Social platforms were not created for long-term case tracking. Therefore, counsellors use creativity and different approaches by, for instance, saving screenshots, logging cases in spreadsheets, writing summaries. It works, but it remains patchy and difficult to standardise. One way forward could be to develop platform-specific evaluation approaches and build closer partnerships with platform providers.

- 3. *Privacy versus proof.*** Anonymity and fleeting interactions are often features of digital interventions leading to counselling in P/CVE. These conditions make it difficult to build relationships and even harder to systematically evaluate impact. In order to help people feel safe when reaching out, practitioners typically need to ask for as little personal data as possible. Conversely, to show their work is making a difference, practitioners need certain information. Balancing both sides means building systems that protect users while still extracting meaningful insights from practitioners' work impact.
- 4. *Mixed-method strategies.*** Limited resources, whether in terms of staff, time, or expertise, pose a major challenge for evaluating digital counselling projects. This makes it difficult to collect data systematically, carry out ongoing qualitative analysis, or further develop evaluation methods. Better resource planning and the use of simple, practical evaluation tools can make a real difference. Easy to use feedback tools, emoji-based check-ins, practitioner journals, and anonymous review mechanisms are just some of the blended methods gaining ground. When clients help define what 'success' looks like, evaluation becomes more about learning together than about ticking boxes, eventually benefitting clients, practitioners, and funders. If appropriate capacities exist, internal measures could be paired with external evaluations to make the overarching approach more robust.

In addition to these core insights, three cross-cutting aspects must be emphasised to ensure more robust and context-sensitive evaluation strategies:

- 1. *Hybrid counselling formats.*** Increasingly, digital counselling does not end online. Many cases blend online interaction with offline follow-up or escalation to further services. Evaluation approaches must account for this spectrum of contacts, moving beyond binary categories of 'digital' vs 'offline' to reflect the hybrid nature of many counselling journeys.
- 2. *Different counselling logics.*** Counselling services follow distinct operational models, whether classic long-term counselling, direct chat-based advice, social media outreach, referral-based triage, or hybrids. Each of these logics comes attached with their own rhythm, risk profile, and evaluative needs. There is no one-size-fits-all model, and evaluation frameworks have to reflect this diversity.
- 3. *Institutional contexts.*** The evaluation framework is also shaped by the implementing organisation's own institutional logic. State-led projects often prioritise accountability and standardised metrics, whereas civil society organisations emphasise flexibility, trust-building, and qualitative insight. Recognising these differences is key to designing fair and realistic evaluation frameworks.

Building on the insights outlined above, we propose the following recommendations for practitioners and evaluators seeking to strengthen their digital counselling strategies and evaluation frameworks:

- ***Institutionalise multi-method evaluation.*** Adopt flexible evaluation frameworks that define and capture unplanned but significant outcomes. Combine ‘hard’ numbers, such as how many people were reached, with ‘soft’ real stories and conversation snapshots. Together, these methods are able to provide a more comprehensive story and enable honing in on ‘impact’ both from both a quantitative and qualitative standpoint.
- ***Invest in interoperable, privacy-by-design tools.*** Implement simple yet systematic monitoring tools to manage resource constraints effectively. Purpose-built dashboards that anonymise data at the source but allow aggregated analysis can ease the documentation burden and enable cross-project benchmarking without violating confidentiality. Build platforms that protect data but still allow teams to compare outcomes and spot trends without compromising privacy.
- ***Cultivate a learning culture.*** Do not perceive evaluation as a favour to funders, but as something that helps strengthen practice. Peer reviews, team reflections, and co-created indicators can turn insights into action.
- ***Bridge the language gap.*** Engage with funders proactively to establish shared language and realistic expectations. Communication between funders, evaluators, and practitioners is often hindered by a lack of shared understanding around goals and impact criteria. To make evaluation processes more effective, the need for ongoing dialogue and the development of a common language to describe the effects of digital interventions should be addressed. Everyone involved, from funders to frontline practitioners, needs shared terms and common goals. Clear categories and transparent mechanisms for sharing what has been learned can help raise standards across the board.

In short: Evaluating digital counselling is not impossible, nor do practitioners reject it, but it does require innovative thinking. By using mixed methods, focusing on safety, and encouraging honest dialogue across teams and multiple sectors, evaluation can be turned into a true driver of support and change, instead of becoming a burden and a hindering factor. Done correctly, it can help improve practice, build trust, and even change lives.

Focus Groups and Expert Interviews

Focus Groups

FGG1: Online focus group held on 15th October 2024, comprising six expert practitioners from classic digital counselling settings.

FGG2: Online focus group held on 21st October 2024, comprising five expert practitioners from outreach counselling work.

Expert Interviews

German Experts

EXP01: Online interview conducted by the authors, 10th March 2024

EXP02: Online interview conducted by the authors, 30th March 2024

EXP03: Online interview conducted by the authors, 2nd May 2024

EXP04: Online interview conducted by the authors, 2nd May 2024

EXP05: Online interview conducted by the authors, 7th May 2024

EXP06: Online interview conducted by the authors, 7th May 2024

In addition, several informal interviews and background conversations were conducted by the authors between March and October 2025.

International Experts

INT01: Online interview conducted by the authors, 15th October 2024

INT02: Online interview conducted by the authors, 31st October 2024

Literature

- Abrams, Zara 2025: Using generic AI chatbots for mental health support: A dangerous trend (American Psychological Association Services, Inc.), 12.03.2025, <https://www.apaservices.org/practice/business/technology/artificial-intelligence-chatbots-therapists> (Last accessed: November 18, 2025).
- Baaken, Till/Becker, Reiner/Bjorgo, Tore/Kiefer, Michael/Korn, Judy/Mücke, Thomas/Ruf, Maximilian/Walkenhorst, Dennis 2018: Herausforderung Deradikalisierung. Einsichten aus Wissenschaft und Praxis. PRIF Report 9/2018, Frankfurt (Main). https://www.prif.org/fileadmin/Daten/Publicationen/Prif_Reports/2018/prif0918.pdf (Last accessed: November 23, 2025).
- Caplan, Gerald 1964: Principles of Preventive Psychiatry, New York.
- Davey, Jacob/Birdwell, Jonathan/Skellett, Rebecca 2018: Counter Conversations: A Model for Direct Engagement with Individuals Showing Signs of Radicalisation Online (Institute for Strategic Dialogue, Technical Report), <https://www.isdglobal.org/isd-publications/counter-conversations-a-model-for-direct-engagement-with-individuals-showing-signs-of-radicalisation-online/> (Last accessed: November 19, 2025).
- ESA (Entertainment Software Association) 2025: Power of Play. 2025 Global Video Games Report, <https://www.theesa.com/wp-content/uploads/2025/09/PoP-2025-v10-web-spreads.pdf> (Last accessed: November 19, 2025).
- Gehrmann, Hans-Joachim 2010: Onlineberatung – eine Herausforderung für die Soziale Arbeit, in: Cleppien, Georg/Lerche, Ulrike (Eds.) 2010: Soziale Arbeit und Medien, Wiesbaden, 105–115.
- Gheorghe, Ruxandra Mihaela/Tarshis, Sandra/Asakura, Kenta 2024: Clinician Perceptions of Online Therapy During the COVID-19 Pandemic: A Qualitative Study, in: Families in Society 105: 4, 716–729.
- Glaser, Barney G./Strauss, Anselm L. 1967: The Discovery of Grounded Theory: Strategies for Qualitative Research, Chicago.
- Kanitz, Juliane/Imhof, Willi/Konradi, Moritz 2021: Evaluation Streetwork@online. Abschlussbericht, Berlin: Camino Werkstatt für Fortbildung, Praxisbegleitung und Forschung im sozialen Bereich GGmbH, https://camino-werkstatt.de/downloads/Abschlussbericht_Eval_streetwork@online.pdf (Last accessed: November 19, 2025).
- Koynova, Svetla 2025: Klient*innenzentrierte Evaluation (in press).
- Khairat, Saif Sherif/Dukkipati, Aniesha/Lauria, Heather Alico/Bice, Thomas/Travers, Debbie/Carson, Shannon S. 2018: The Impact of Visualization Dashboards on Quality of Care and Clinician Satisfaction: Integrative Literature Review, in: JMIR Human Factors 31: 5 (2), e22.
- Leong, Frederick T. L. 2008: Toward a Global Vision of Counselling Psychology., in Counselling Psychology, edited by Frederick T. L. Leong and Mark M. Leach, 115–32. Abingdon: Routledge.
- Limbic AI 2025: Clinical AI for Mental Healthcare, <https://www.limbic.ai/> (Last accessed: November 19, 2025).
- Kellimni n.d.: Support is one click away. Find your voice, Find your strength (Kellimni.com – Digital Youth Counselling Platform), <https://kellimni.com/> (Accessed: November 19, 2025).
- Molenkamp, Merel/Wouterse, Lieke/Gielen, Amy-Jane 2018: Ex Post Paper. Guideline Evaluation Of PCVE Programmes And Interventions (RAN MS Workshops On PCVE Evaluation, July 2018), https://home-affairs.ec.europa.eu/system/files/2020-09/ms_workshops_guidelines_evaluation.pdf

- tion_of_pcve_programmes_and_interventions_july_2018_en.pdf (Last accessed: November 19, 2025).
- Moonshot 2025: Moonshot CVE – Targeted Interventions Against Online Extremism, <https://moonshotteam.com/> (Last accessed: November 19, 2025).
- Ni, Yang/ Jia, Fanli 2025: A Scoping Review of AI-Driven Digital Interventions in Mental Health Care: Mapping Applications Across Screening, Support, Monitoring, Prevention, and Clinical Education, in: *Healthcare* 13: 10, 1–23.
- OSCE 2019: Understanding Referral Mechanisms in Preventing and Countering Violent Extremism and Radicalization That Lead to Terrorism. Navigating Challenges and Protecting Human Rights. A Guidebook for South-Eastern Europe, Vienna: Transnational Threats Department, OSCE Secretariat, <https://www.osce.org/files/f/documents/7/4/418274.pdf> (Last accessed: November 19, 2025).
- LAH (Life After Hate) 2024: Life After Hate Releases Discord Channel, Engaging with Today’s Violent Extremists Where They Digitally Live (The Daily Former), <https://www.lifeafterhate.org/blog/2024/5/23/the-daily-former-discord/> (Last accessed: November 19, 2025).
- Thiery, Heinz 2018: Was ist Online-Beratung? Ein differenztheoretischer Vergleich (Deutschsprachige Gesellschaft für psychosoziale Onlineberatung (DGOB), Vorstandsmaterialien Nr. 1/ 2018), <https://dg-onlineberatung.de/wp-content/uploads/2022/09/Definition-Online-Beratung-Langtext.pdf> (Last accessed: November 19, 2025).
- Urman, Aleksandra/Katz, Stefan 2020: What They Do in the Shadows: Examining the Far-Right Networks on Telegram., in: *Information, Communication & Society* 25: 7, 904–923.
- Van de Donk, Maarten/Ruf, Maximilian 2022: Exploring hybrid and digital rehab work (Conclusion Paper, RAN REHAB Working Group Meeting), https://home-affairs.ec.europa.eu/system/files/2022-08/ran_rehab_exploring_hybrid_digital_work_24-25032022_en.pdf (Last accessed: November 19, 2025).

Appendix

To further highlight the fact that all practitioners who participated in this study share a keen interest in meaningful MEL processes, and to provide inspiration and first guidance to other practitioners, evaluators, and funders in charge of future evaluations in this field, the authors compiled a first list of potential indicators that can be used to evaluate digital counselling. These are solely based on the focus group discussions conducted for this report and, in line with the general findings of this text, they should not be viewed as a final and comprehensive set of indicators, but rather as guidance and inspiration on the path towards developing individual, project-specific indicators for success.

Quantitative Indicators	
Indicator	Source
Number of individual chats/interactions	FGG2, Pos. 88, 90
Number of publicly visible interactions (e.g., comments, likes, shares)	FGG2, Pos. 88
Length of interaction (e.g., minutes, length of conversation)	FGG2, Pos. 88
Group/channel/etc. membership numbers	FGG2, Pos. 78
Number of content clicks, views, likes (on <i>TikTok</i> , <i>Twitch</i> etc.)	FGG1, Pos. 57; FGG2, Pos. 88
Number of contacts to be achieved	FGG2, Pos. 96
Number of filled-in post-counselling surveys	FGG2, Pos. 84
Number of workshops/events	FGG1, Pos. 63
Number of participants during events	FGG1, Pos. 63
Time on site / dwell time on websites	FGG1, Pos. 57
Number of publication downloads	FGG1, Pos. 63
Project acceptance by interested third parties (e.g., via invitation to participate in live streams)	FGG1, Pos. 63
Number of counselling cases per year	FGG1, Pos. 18

Qualitative Indicators	
Indicator	Source
Content and depth of interaction/conversations	FGG2, Pos. 80
Assessment of radicalisation risks (e.g., through ideological references)	FGG2, Pos. 80
Gradual changes throughout the course of conversation/interaction	FGG2, Pos. 80
Changes in the belief system or small changes of attitudes	FGG2, Pos. 80
Ability to self-reflect	FGG2, Pos. 80
Developments and experiences of (potential) clients	FGG2, Pos. 80
Observable needs of clients	FGG2, Pos. 94
Recurring (themes of) needs	FGG2, Pos. 94
Trends and topics that work well in counselling situations	FGG2, Pos. 94
Quality of (line of) arguments in digital debates	FGG2, Pos. 74
(Analysis of) interruption of conversations and relationships	FGG2, Pos. 100
Perceived impact of support services	FGG2, Pos. 54
Successful referral to other counselling/support services	FGG2, Pos. 90
Participant and partner feedback	FGG1, Pos. 46
Existence and quality of self-evaluation and formative methods of concept development	FGG1, Pos. 46, 48, 50, 60
Regular introspection and reflection of working modalities	FGG1, Pos. 48
Exchange with external partners (e.g., from research, civil society)	FGG1, Pos. 50

Qualitative Indicators	
Topical debate and trend monitoring in relation to project objectives (e.g., on gaming, right-wing extremism online)	FGG1, Pos. 52
Presence of intervision, meta-reviews, objective hermeneutics, or similar approaches	FGG1, Pos. 54
Qualitative exchanges with funders beyond numerical reporting	FGG1, Pos. 66, 67
Use of qualitative case reports/summaries	FGG1, Pos. 67

Trust, Tone, Timing, and Tech. Assessing Digital P/CVE Counselling Services and Challenges for Evaluation Practice

Juliane Kanitz | Maximilian Campos Ruf |
Svetla Koynova | David Tschöp | Franziska Heil

With funding from the



by decision of the
German Bundestag

Digital technologies are increasingly finding their way into individual counseling work in secondary and tertiary prevention. In addition to traditional telecounseling, interactive platforms, chatbots for initial contact and even the first AI-supported applications are now being used. These developments open up opportunities, but also pose new challenges for specialist practice and evaluation research: How can the quality and impact of such digitally supported services be appropriately assessed? And what methodological and ethical questions arise?

This *PrEval* pilot study looks at current developments in Germany, Europe and selected international contexts. Based on qualitative surveys such as interviews and focus group discussions in the period from 2024 to 2025, it maps key challenges, collects findings from initial practical approaches and develops practical suggestions on how evaluations can be designed in this dynamic field.

The results provide valuable starting points for evaluators, political decision-makers, funding providers and academia to further develop the quality management of digital counseling services.

PrEval

Zukunftswerkstätten



Peace Research Institute Frankfurt
Leibniz-Institut für
Friedens- und Konfliktforschung