Visualizing Self-Report Data for Clinical Insight

Practitioner Perspectives on ESM Feedback for Assessing Therapy Effectiveness

Jules Savelkoul[1]

Supervisor(s): Dr. ir. Willem-Paul Brinkman, Esra de Groot

EEMCS, Delft University of Technology, The Netherlands

A Thesis Submitted to EEMCS Faculty Delft University of Technology,
In Partial Fulfilment of the Requirements
For the Bachelor of Computer Science and Engineering
June 22, 2025

Name of the student: Jules Savelkoul
Final project course: CSE3000 Research Project
Thesis committee: Dr. ir. Willem-Paul Brinkman, Esra de Groot, Prof.dr.ir. R.L. Lagendijk

Abstract

Background: The Experience Sampling Method (ESM) enables the collection of momentary self-reports on thoughts, emotions, and behaviour in daily life. However, there is limited practical guidance on how to visualize this data to support practitioners.

Objective: This study aimed to explore how ESM data can be visualized to effectively communicate treatment effectiveness to mental health practitioners.

Methodology: A design-based approach was used to guide the development of two ESM visualizations. Visualization 1 consisting of a single temporal line graph and Visualization 2 consisting of a dashboard-style layout with four graphs. Practitioners, psychology students and medicine students completed a questionnaire evaluating the clarity, interpretability, and perceived clinical utility of the visualizations. Thematic analysis was conducted on the responses.

Results: Analysis of the responses from ten participants revealed four themes: (1) preference for single temporal line graphs due to their clarity and intuitive presentation of change over time, (2) usability challenges in complex dashboards, particularly with dual axes and ambiguous terminology, (3) value of contextual variables, such as interactions and homework completion, in understanding therapy progress and (4) a strong desire for customization to fit client-specific needs.

Conclusion: Clear visualizations of ESM data may support practitioners in assessing therapy effectiveness, especially when they enable reasoning about both therapeutic change and contextual influences. However, individual differences in interpretation underscore the need for flexibility and customizability.

1 Introduction

Mental health treatments have increasingly shifted the focus from pure symptom reduction towards a more individual approach, where a client's personal treatment goals are prioritized [20]. Alongside this development, there are strong indications that a better understanding of a client outside of the clinical setting would be complementary [18].

The experience sampling method (ESM) has emerged as a promising approach to achieve this [7; 5; 1; 23]. ESM is a technique that allows real-time monitoring in an individual's natural environment, meaning that thoughts, feelings, and behaviour can be captured at the moment of experiencing [23]. Not only would gathering information in context give mental health practitioners a broader and more accurate view of a client's experiences, but it would also help address the problem of recall bias -that is, the tendency for individuals to misremember or distort past thoughts, feelings, or behaviors when asked to report them retrospectively [24].

Furthermore, ESM may provide the added benefit of monitoring clients' treatment responses, potentially enhancing the subjective insights of practitioners. This is highly relevant, since therapist are often unable to predict therapy outcomes correctly and mostly overestimate the effectiveness of the therapy on their clients [26; 15]. By being able to better identify a client's progress, the possibility to find an optimal treatment plan or intervention for a specific person arises [3].

However, the usefulness of this feedback largely depends on clear presentation of the raw data. Practitioners often express insecurities regarding their interpretation abilities and desire intuitive designs that offer more insight [5].

The majority of visualization tools for ESM data rely on statistical techniques like vector autoregression (VAR), that come with several difficulties when displaying temporal data and fluctuations in data [21]. ESMvis however, an ESM visualization tool build using the R programming language, was developed to allow dynamic visualization of ESM data [7].

While ESMvis enables dynamic visualization of ESM data, its full potential needs further systematic research [7]. Current literature offers little practical guidance on how to present this complex data in a way that is both intuitive and clinically useful [10]. This paper will focus on the visualization of treatment effectiveness, leading to the following research question: **How can ESM data be visualized to effectively communicate treatment effectiveness to mental health practitioners?** Exploring this question begins with developing a solid understanding of ESM and reviewing current research and related work. Building on this foundation, visualization prototypes are designed and then evaluated, preferably by mental health practitioners. The results gathered from their feedback are analyzed thematically to draw meaningful conclusions. The contributions of this paper are fourfold: (1) an analysis of practitioner preferences regarding different graph types and data displays; (2) insight into which aspects of ESM data are considered most informative; (3) design principles grounded in practitioner reasoning; and (4) an improved understanding of which visualizations are preferred and why.

The remainder of this paper is structured as follows: Section 2 reviews the relevant literature on ESM and treatment effectiveness. Section 3 outlines the methodology, including the design of visualization prototypes and the evaluation procedure. Section 4 presents the results of the practitioner feedback. Section 5 addresses responsible research considerations, including ethical approval. Section 6 discusses the findings, limitations, and implications. Section 7 concludes with directions for future work.

2 Related Work

The experience sampling method (ESM) has gained traction as a valuable tool in mental health research, offering a way to monitor individual experiences in real-time and real-world contexts. Recent frameworks like ESMvis have aimed to improve the presentation of this rich, temporally dynamic data by introducing visualizations that show fluctuations over time, rather than relying on aggregated statistical summaries such as means or correlations [7]. This research shows that early user feedback has been positive, but its actual clinical utility remains largely unexplored.

Studies exploring the clinical integration of ESM tools have identified a range of challenges. Practitioners report uncertainty in interpreting complex outputs, such as network graphs, and prefer more conventional representations like time series, bar charts, or pie charts [24]. These visual preferences point to a broader issue: ESM data is often difficult to translate into clinically useful insights without additional guidance. Interviews with therapists have also highlighted the need for clear, intuitive visual designs that take less time to interpret and a better understanding of what information is most helpful to display [12].

Another key issue raised in prior research is the burden ESM may place on clients, particularly in long-term or high-frequency data collection settings [5]. It shows that flexibility and customization, in both ESM questionnaires and information display, are often cited as necessary features, to maintain client engagement and to mitigate potential harm or fatigue. Practitioners in this study also emphasize that ESM tools should complement, rather than replace, traditional therapy approaches.

Efforts to develop feedback tools based on ESM data have also underlined the importance of co-creation with end users [11]. This study evaluated a self-report app designed to provide personalized feedback to adolescents. Despite its promise, the app showed no significant treatment improvements, and engagement was a persistent challenge, reinforcing the need for further work in making such tools both effective and usable.

Overall, these studies suggest that while ESM has great potential for enhancing mental health care, its utility depends heavily on how the data is presented and interpreted. There remains a notable gap in research on which visual formats best communicate useful insights to mental health practitioners. This study addresses that gap by exploring practitioners' preferences and reasoning around various visual representations of ESM data.

3 Methodology

To explore practitioner preferences and reasoning regarding ESM data visualizations, this study adopted a design-based approach: repeatedly designing, evaluating, and refining prototypes based on practitioner feedback. This paper reports one such design cycle. Instead of developing a fully functional visualization tool, a set of mock-up designs was created with the aim of eliciting rich, qualitative feedback from mental health practitioners. This choice was motivated by two key considerations: first, practitioners must find visualizations not only clear but also intuitively useful; second, it is premature to invest in fully implemented systems without a grounded understanding of what kind of information should be visualized and how it is best presented [13]. Equally important is designing in close collaboration with practitioners [12], as they are the end users who must find the visualizations both intuitive and clinically useful.

The designs were intentionally varied and included both well-argued elements as well as elements that were found to be ambiguous. This was done to stimulate feedback explaining the underlying reasons why certain design choices were seen as more or less useful. Understanding these reasons is essential for future research and developing tools for ESM that actually support clinical work.

The visualizations focused on showing several types of information: (1) client progress, (2) client adherence, and (3) context. Only showing progress is not sufficient to show effectiveness, as treatments are never identical [16]. Context can be very valuable to display and adherence to treatment (e.g. whether out-of-session tasks are performed) as well. Without clients putting in the effort, it cannot be expected to see optimal results and thus showing only progress does not give a full picture [14].

3.1 Visualization Design

The mock-up visualizations were developed using Figma, a collaborative web-based interface design tool. Although actual patient data were not required, visualizations were inspired by an ESM questionnaire framework used in mindfulness-based therapy [4], which tracked, among other variables, whether and when mindfulness exercises were completed. In the designs presented here, these completed tasks serve as a proxy for any type of out-of-session therapeutic exercise. Furthermore, it should be noted that the designs present ESM data of a single client.

Client progress was conceptualized in alignment with Routine Outcome Monitoring (ROM) practices, typically measured through repeated assessment of mental health symptoms and perceived well-being. ROM has demonstrated to improve treatment outcomes [2; 16].

When considering which contextual information to include in the visualizations, the challenge was to ensure clinical usefulness. Only information that meaningfully contributes to the effectiveness of treatment should be displayed [19]; otherwise, it risks creating unnecessary visual clutter. Determining which variables are truly useful is complex, because each client's situation is unique and their goals, symptoms, and relevant contextual factors differ [16]. This variability suggests that allowing customization of contextual data in the visualizations could help practitioners focus on the factors most relevant to individual clients and therapeutic approaches.

Practitioners generally prefer visualizations that provide concrete insights into changes over time [2]. Consistent with this preference, several studies indicate that longitudinal line graphs are effective for depicting symptom progression or behavioral changes [8; 7; 6]. Furthermore, the ability to focus on specific symptoms or behaviours through customization enhances clinical utility of these visualizations [8]. In contrast, bar graphs have been found valuable for delivering quick overviews or summaries of client states and behaviors over defined periods [6]. Combining both visualization types offers complementary benefits: line graphs capture temporal dynamics, while bar graphs provide straightforward snapshots. Given the necessity for clarity and ease of interpretation for effective clinical use [24], the designs developed for this study incorporate a combination of line and bar graphs. This dual approach aims to meet practitioners' needs for both

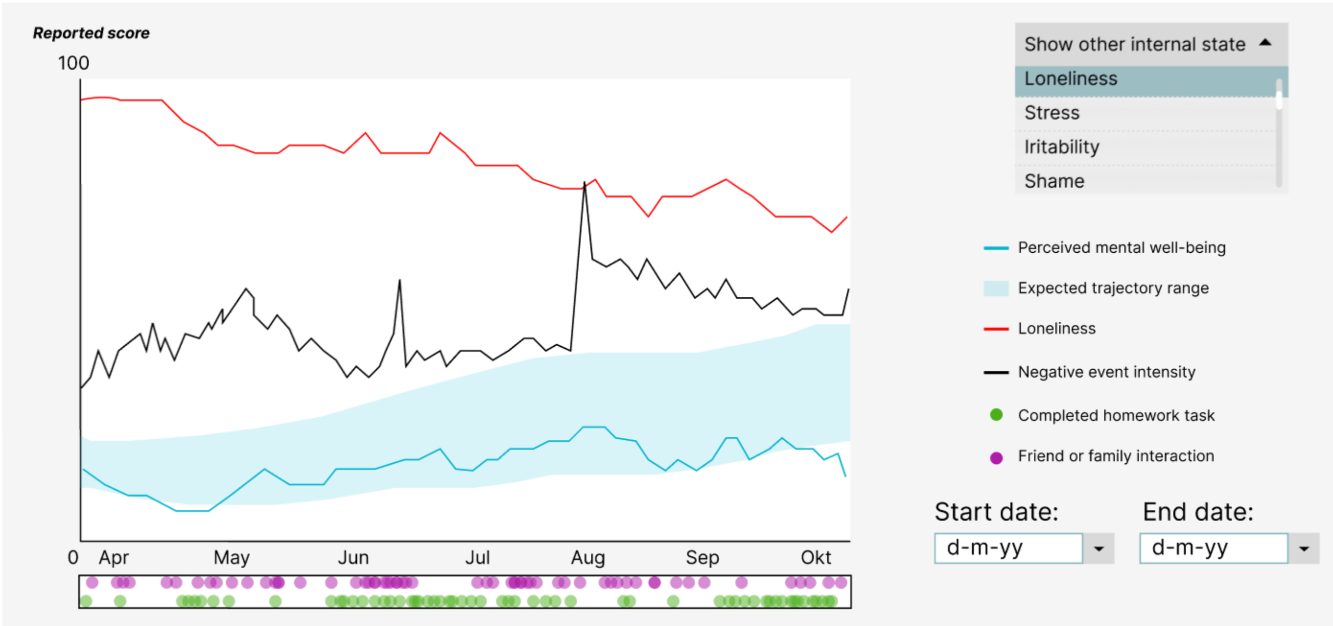

Figure 1: Visualization 1 presenting mocked ESM data of a single client over roughly six months

detailed progress tracking and rapid assessment.

Finally, some studies suggest that integrating expected symptom progress trajectories into visualizations is both feasible and helpful [12; 2]. They show that displaying normative or anticipated recovery patterns can aid clinicians in identifying clients who are not progressing as expected, thus enabling timely intervention adjustments.

3.2 Visualization Principles

In order to explore how different design choices impact understanding and perceived usefulness of ESM data visualizations, two mock-up prototypes were developed. The goal was to examine how clarity, intuitiveness, and perceived clinical utility were affected by factors such as graph type, number of elements presented, and level of contextual detail. Visualization 1 presents a single temporal line graph graph combining several variables in one view. This approach emphasizes the ability to recognize patterns and fluctuations over time. However, displaying multiple variables in one figure could risk visual overload. Visualization 2 uses a dashboard layout consisting of four separate graphs, using both line and bar charts. This design allows to compare visual representation (e.g., fluctuation vs. endpoint summaries) and to choose the most intuitive format. However, segmenting information and using possibly illogical graphs may reduce perceived clinincal utitilty and causes misunderstanding. This comparative setup aims to investigate whether integrating multiple data streams into a single graphs improves or hampers clarity, which variables are deemed useful, how important visible fluctuations are in assessing client response and progress, and whether multiple complementary visualizations compromises usability.

Visualization 1: Single Longitudinal Line Graph

Design 1 displays a longitudinal line graph in which several elements are visualized based on the above information (see Figure 1). It shows a general mental well-being line, one chosen internal state and perceived intensity of negative events. Alongside the client's progress, it also shows an expected trajectory range, since this would be possible [12] and is also useful for detecting clients that are not on track [2]. However, this conflict with the idea of treatments being vastly different and incomparable. Using this specific element, therefore is only shown in visualization 1 and not in visualization 2.

Additionally, there are two other contextual features shown in the design, which could result in the visualization being perceived as overwhelming. Underneath the line graph, a practitioner can view the moments in time when their client has had meaningful social interaction and where an out-of-session task was completed.

Visualization 2: Dashboard Consisting of Four Different Graphs

Visualization 2, displays a dashboard of 4 separate graphs (see Figure 2). The upper left and right graphs show the same elements in a different way. Changes are separated into emotional, physical, and behavioural changes, in order to create a clearer and more organised overview. The upper left shows a temporal graph with fluctuations over time, whereas the upper right shows a bar graph depicting the change from start to end date without fluctuations. Since preferences among practitioners on visualization possibilities vary vastly, showing both could give the practitioner the choice to read the graph that is most insightful or intuitive to them [24; 12].

The lower left graph shows a bar chart explaining the daily average amount of tasks completed, together with the

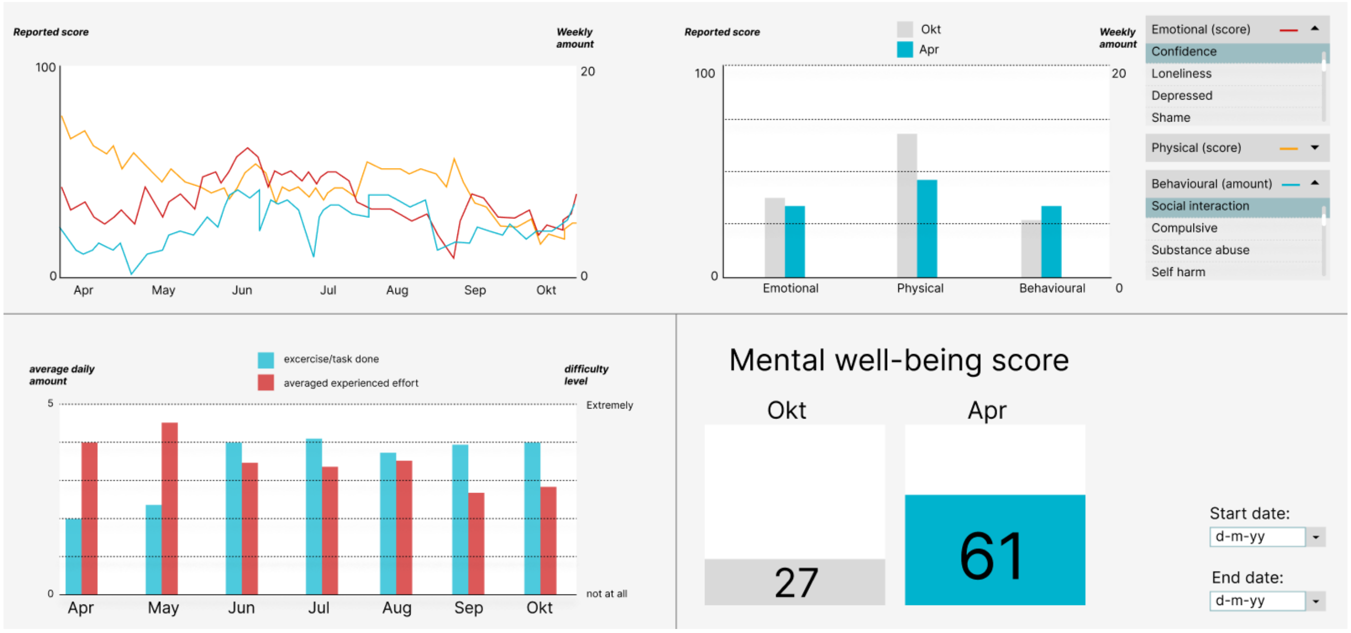

Figure 2: Visualization 2 presenting mocked ESM data of a single client over roughly six months

perceived effort, making it a mix of both subjective and objective measurements. Lastly, the lower right gives a quick overview of the difference in the mental well-being score.

3.3 Questionnaire Design and Participants

Feedback was collected through an online survey made with *Qualtrics* [22] that included open-ended questions (see Table 1). These questionnaires allowed for scalable data collection and rich thematic insights into practitioner reasoning. Questions focused on interpretability, usefulness, and design preferences, with a special emphasis on the *why* behind evaluations. Two of the three questions were asked twice: once for Visualization 1 and once for Visualization 2. To minimize order effects, the presentation order of the visualizations was randomized.

Due to high workload constraints among practicing clinicians [25], the initial idea was to distribute the questionnaire to mental health practitioners and psychology students. While not yet fully qualified professionals, these participants were assumed to be able to evaluate the designs from a clinical perspective. However, collecting a sufficient amount of responses took more time than anticipated. Therefore, the same questionnaire was distributed to medicine students as well. In addition to the questions shown in Table 1, the questionnaire collected information about the background, age and gender of the participant.

3.4 Data Analysis

Qualitative responses from open-ended questions were analyzed using thematic analysis, following Braun and Clarke's six-step method [9]. First, all responses were read thoroughly to become familiar with the data and to gain an overall understanding of emerging patterns. Initial coding was then performed using an inductive approach, allowing for open coding that captures salient features of the data without being restricted to pre-existing categories. While some preliminary themes were identified based on the research questions and relevant literature, the coding process remained flexible, allowing for modification or addition of themes as new insights emerged.

To enhance the reliability and validity of the analysis, a second independent coder was engaged to perform double coding on a subset of the data. Through an inter-coder reliability measured with Cohen's Kappa [17], every code got a score. Codes with a weak level of agreement (below 0.60) underwent a review. These codes were reexamined, and either merged, refined or excluded to enhance consistency and reliability.

4 Results

A total of ten individuals participated in the study. Of these, nine provided comprehensive responses to all five survey questions. One participant completed four of the five questions, albeit with limited detail. Despite the brevity of this participant's contribution, their response was retained in the analysis to ensure the inclusion of all available data. Demographic data revealed that most participants were aged between 21 and 30 years, and a substantial majority reported no prior familiarity with Experience Sampling Method (ESM) (see Table 2). The participant cohort included not only mental health practitioners and psychology students but also a few medicine students. Their inclusion enriches the study by offering perspectives on visualization design that are relevant across a broader spectrum of clinical healthcare disciplines.

After getting familiar with the collected responses, the coding assessment was applied. Having two independent coders, the inter-coder reliability could be calculated using Cohen's

Question	Goal	Visualization
How do you go about reading or interpreting this visualization? Describe your thought process. While doing this, explain why certain elements are confusing or clear.	Interpretation approach; identify clarity issues.	1 and 2 separately
What elements of this visualization do you believe to be valuable for a better understanding of this client's response to the treatment? Why?	Value assessment; focus on usefulness.	1 and 2 separately
Looking at both visualizations, which elements would you keep/add/remove/change and why?	Comparative feedback; design suggestions.	1 and 2 together

Table 1: Questions for User Evaluation

Kappa. Initially, the individual code reliability values ranged from 0.342 (minimal) to 1.000 (perfect). Where necessary, codes where reexamined. Across 62 codes from five questions, the average coefficient calculated was 0.862, indicating strong agreement [17].

Following this coding assessment, four overarching themes emerged from the data: *Perceived Strengths of Single Temporal Line Visualizations, Usability Challenges and Design Considerations, Perceived Clinical Utility of Visualized Variables, Desire for Customization and Interactive Control* (see Figure 3).

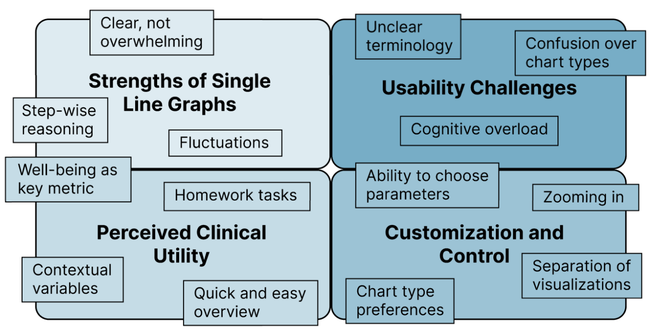

Figure 3: Overview of the four discussed themes together with most important findings

Strengths of Single Temporal Visualizations

Most participants preferred Visualization 1, citing its clarity and its ability to present information without overwhelming the viewer. Several participants (n=6) described using a step-wise reasoning process, scanning the timeline for patterns and relationships between variables. One participant remarked: "With the second peak, you can also see that it affected the completed homework tasks". Almost everyone noted that they could clearly observe (slight) improvements in mental health indicators over time in this visualization. Line graphs were especially appreciated for their ability to depict fluctuations over time. One participant mentioned they found the large legend helpful in understanding the overlay of variables Visualization 1. Another participant had no difficulties interpreting Visualization 1, but found the multiple graphs and double y-axis, used in top-left line graph, confusing. A few participants (n=2) noted that while summary statistics, like those in the top-right panel of Visual 2, were informative, they lacked sufficient nuance: "That would be too black and white." One participant misinterpreted a trend in Visual 1, mistakenly concluding that loneliness increased, suggesting the need for added guidance or annotations.

Usability Challenges and Design Considerations

Terminology was occasionally unclear across both visualizations. Terms like "negative event intensity" and interactive prompts such as "show other internal state" that are both used in Visualization 1, led to confusion (n=2). The lower left graph in Visualization 2 was received especially poorly, with five participants (n=5) stating they could not determine what it represented or how it contributed to a better understanding of treatment effectiveness. Only one participant found it interesting, whereas others described it as "confusing" and "illogical." The placement of the legend in the top-right of Visualization 2 was also problematic for participants (n=2), with one of them missing it entirely. Interestingly, some participants (n=3) who found Visualization 1 clear still experienced cognitive overload when interpreting the line graph in Visualization 2. They expressed confusion on the display of multiple graphs and the double y-axis.

Perceived Clinical Utility of Visualized Variables

The majority of the participants (n=8) expressed strong support for including perceived mental well-being as a core metric, especially when contextualized by other variables. As one participant stated, "...various parameters can provide a broader picture and offer an explanation." Variables such as homework completion, interactions, and negative events were all mentioned as clinically useful when shown in relation to well-being scores. The breakdown of mental health into emotional, physical, and behavioral dimensions, as featured in Visualization 2, was also well-received by several participants (n=3). They felt this distinction could help focus therapeutic discussions on areas of difficulty and even guide treatment planning. One participant noted: "It gives direction to the therapy." There was also appreciation for the bottom-right comparison graph in Visualization 2, which simply depicted two points in time. Participants (n=3) described this as an effective snapshot of progress, offering a quick and easy interpretation, especially in time-constrained settings. The "expected trajectory" line received mixed feedback: one participant valued it for tracking client progress, while another argued it was too prescriptive, noting that progress is highly individual and difficult to standardize.

Desire for Customization and Interactive Control

The need for customizability emerged as the final theme. Multiple participants (n=6) stressed that the relevance of variables depends on the client's situation. While not all explicitly mentioned customization features, several suggested the importance of being able to choose parameters, filter data, or focus on specific periods. One participant explained that separating the visualizations would help them "interpret each part individually and decide what to include in their assessment." Preferences for chart types also varied: while some favored line graphs, others preferred bar charts for certain metrics. One participant explicitly suggested to only keep the line graph in visualization 2. An additional feature suggested was the ability to zoom in on specific time periods, such as a particular month, to enable more detailed clinical discussions around specific events.

	Participants (N=10)	
	No.	**%**
Gender		
Man	2	20
Woman	7	70
Other	1	10
Age		
21-30	9	90
31-50	0	0
51+	1	10
Background		
Mental Health Practitioner	3	30
Psychology Student	4	40
Medicine Student	3	30
ESM Familiarity		
None	9	90
Some	1	10

Table 2: Participant Demographics

5 Discussion

This study aimed to explore how ESM data can be visualized to support mental health practitioners in assessing therapy effectiveness. A thematic analysis revealed four key themes, pointing to the central role of clarity and interpretability in perceived clinical utility. However, differing interpretation skills and preferences in graph types suggest that no single format will work equally well for all users. The following sections interpret these findings and relate them to existing literature.

5.1 Interpretation of Results

Participants largely preferred Visualization 1, a single temporal line graph due to its intuitive display of fluctuations over time. This preference aligns with earlier work emphasizing the value of time series for showing dynamic patterns [6; 24]. These kinds of visualizations appear to support practitioners in evaluating progress and identifying treatment-relevant patterns.

Possibly through this format, participants were able to frequently engage in relational reasoning, linking changes across variables, such as how mood patterns related to out-of-session task completion. This suggests that ESM visualizations can indeed facilitate meaningful clinical insight when multiple variables are shown in a well-structured and coherent manner.

In contrast, the dashboard-style layout in Visualization 2 was perceived as confusing by several participants, particularly due to unfamiliar graph types and dual y-axes. This supports previous findings that conventional and clearly structured visuals are typically preferred [24]. While the goal of dashboards is to offer a comprehensive overview, these results highlight a potential for cognitive overload. This finding differs somewhat from positive responses to dashboard-like formats such as ESMvis [7], suggesting that layout effectiveness likely depends on both design quality and user familiarity. It also points to variability among practitioners in their ability and willingness to interpret more complex displays.

Confusion and uncertainty were further reflected in questions and misinterpretations during the evaluation. While familiarity over time might improve understanding, this places an additional burden on practitioners. Though participants did not explicitly mention a need for user training or instructional support, as found in other studies [5; 8], such guidance could enhance clinical utility without significantly increasing time investment.

Participants also appreciated the inclusion of contextual variables. Especially impact of client adherence, meaningful interactions and negative events were mentioned. These were seen as helpful in interpreting changes in well-being, suggesting a possible preference for complete overviews. The breakdown of mental health into emotional, physical, and behavioral dimensions was viewed by some as useful for guidance in therapy. These findings suggest that being able to tailor variables to individual clients may support a more nuanced understanding of a client's response to treatment. This aligns with research showing that practitioners vary in how they use ESM data, focusing on the most relevant aspects for their specific client [3].

This need for tailoring was reflected in the strong desire for flexibility and customization. Several participants emphasized the importance of filtering data, zooming into specific timeframes, and selecting relevan variables. Such preferences are in line with literature advocating for co-design and adaptive tools to accommodate diverse clinical needs [11].

These visualization preferences, particularly for clear temporal representations, inclusion of contextual variables, and the ability to focus on specific timeframes, are especially relevant for assessing therapy effectiveness. They support practitioners in tracking change, reasoning about intervention effects, and connecting clinical progress to situational context.

5.2 Limitations

Although the present study provides information on why certain elements of visualizations are desired, several limitations must be considered. First, this study focused exclusively on the perspectives of practitioners. The clients' view on visualization utility were not included. Second, participants were not provided with detailed explanations of the ESM methodology or its potential advantages, which may have limited their appreciation of the visualizations' value. Third, the study focused on interpretability and perceived usefulness, rather than evaluating whether visualizations actually improve decision-making or treatment outcomes. Fourth, the sample included both practicing professionals and students, but the small sample size limits generalizability and may not fully represent the diversity of practitioner experiences. Finally, the use of task completion as a proxy for therapeutic engagement may have contributed to mixed responses. Its meaning was not clearly defined, which may have reduced its interpretability and perceived relevance.

6 Responsible Research

Responsible research practices were carefully considered throughout the study to ensure ethical integrity, protect participant rights, and promote transparency. This included attention to ethical approval procedures, informed participation, secure data handling, and thoughtful recruitment practices.

Ethical Approval and Informed Consent

Prior to the start of the study, approval was obtained from the Human Research Ethics Committee (HREC) of TU Delft (ID = 5405). All research procedures complied with the ethical standards set by the institution. Participants were provided with comprehensive information about the study's purpose, procedures, and expected duration. An informed consent form was placed at the beginning of the questionnaire, and participants could not proceed without explicitly agreeing to the terms. This ensured that all participants voluntarily consented to the collection and use of their data.

Data Privacy and Use of External Tools

To safeguard participant privacy, the option to collect IP addresses and other identifying metadata was disabled within Qualtrics. Furthermore, only aggregated and anonymized analysis results are reported. In cases where individual quotes are used to support thematic insights, identifying information is removed. While this approach aligns with ethical research practices and enhances confidentiality, it may limit full reproducibility of the study. It is also worth noting that while Qualtrics complies with international data protection standards, using a third-party platform introduces a degree of reliance on external data infrastructure. Although minimal, this technical consideration may carry a small risk regarding data security.

Ethical Considerations in Real-World Use

As ESM-based tools move toward clinical integration, several ethical considerations must be addressed to ensure responsible deployment. First, visualizations based on sensitive patient data introduce inherent privacy risks. While this study used anonymized mock-ups and did not collect real patient data, any future implementation must comply with data protection regulations and secure infrastructures to prevent unauthorized access or misuse.

Second, there is a risk that practitioners may come to overrely on visualized ESM data when evaluating client progress. Although such data can provide valuable real-time insight, it may not capture the full complexity of a person's mental health state. Relying too heavily on ESM feedback could result in oversimplified or misguided treatment decisions, particularly if used without narrative information from therapy sessions.

Finally, these tools may inadvertently shift clinical practices by replacing rather than complementing current assessment methods. While digital support systems can enhance care, their introduction should be guided by rigorous evaluation and active collaboration with practitioners to ensure they serve as a complementary tool and not substitute existing therapy approaches.

7 Conclusion

This study explored how ESM data can be visualized to support practitioners in evaluating therapy effectiveness. The findings underscore the importance of clear, interpretable visualizations for fostering clinical insight. However, differences in preferences and interpretation skills among practitioners suggest that no single visualization will be optimal for all users. Participants valued the inclusion of contextual and well-being related variables, as these were seen to influence clinical progress. Temporal line graphs with multiple variables allowed for this relational reasoning, possibly supporting a clearer assessment. In addition, they emphasized the need for flexibility in exploring and customizing data displays. More complex, dashboard-style layouts may allow for complementing graphs that offer additional insight. However, they introduced challenges in this study related to cognitive load and clarity. These results point to the importance of designing visualization tools that are both intuitive and adaptable to user needs and contexts, as not only practitioners' skills vary, but clients' experiences and trajectories as well.

7.1 Future Work

Future research should involve more in-depth interviews to better understand how practitioners interpret and apply ESM visualizations in practice. Moreover, studies should move beyond perceived usefulness and examine how visualization tools affect clinical decision-making and treatment outcomes in practice. Further co-creation efforts with practitioners, and maybe clients, remain valuable and should include higher-fidelity, interactive prototypes as well. Testing these designs with guidance, either system-based or supported by a facilitator, could help interpretation gaps. Finally, given the diverse needs of practitioners and clients alike, tailoring ESM tools to specific therapeutic contexts remains essential.

8 Acknowledgements

Special thanks to Esra de Groot and Dr. Ir. Wilem-Paul Brinkman for their guidance in conducting the research and feedback on earlier stages of the paper, and Cristiana Cotoi for serving as a second coder. Moreover, AI was used to a small extent to structure thoughts or ideas when overview was lost.

References

[1] Yasmeen Abdullah Ali Anjeer Alshehhi, Khlood Ahmad, Mohamed Abdelrazek, and Alessio Bonti. Proceedings of the 35th australian computer-human interaction conference. In *Data-Task-Visualisation Analysis based on Top mHealth Apps*, 2023.

[2] Michael Barkham, Kim de Jong, Jaime Delgadillo, and Wolfgang Lutz. Routine outcome monitoring (rom) and feedback: Research review and recommendations. *Psychotherapy Research*, 2023.

[3] Jojanneke A. Bastiaansen, Yoram K. Kunkels, Frank J. Blaauw, Steven M. Boker, Eva Ceulemans, Meng Chen, and et al. Sy-Miin Chow. Time to get personal? the impact of researchers choices on the selection of treatment targets using the experience sampling methodology. *Journal of Psychosomatic Research*, (1), 2020.

[4] Teresa Bolzenkötter, Andreas B. Neubauer, and Peter Koval. Impact of a momentary mindfulness intervention on rumination, negative affect, and their dynamics in daily life. *Affective Science*, 2025.

[5] Fionneke M. Bos, Evelien Snippe, Richard Bruggeman, Marieke Wichers, and Lian van der Krieke. Insights of patients and clinicians on the promise of the experience sampling method for psychiatric care. *PS*, 2019.

[6] Fionneke M Bos, Lino von Klipstein, Ando C Emerencia, Erwin Veermans, Tom Verhage, Evelien Snippe, Bennard Doornbos, Grietje Hadders-Prins, Marieke Wichers, and Hariëtte Riese. A web-based application for personalized ecological momentary assessment in psychiatric care: User-centered development of the petra application. *JMIR Mental Health*, 2022.

[7] Laura F. Bringmann, Date C. van der Veen, Marieke Wichers, Hariëtte Riese, and Gert Stulp. Esmvis: a tool for visualizing individual experience sampling method (esm) data. *Qual Life Res*, 2021.

[8] Sarah Chang, Lucy Gray, Noy Alon, and John Torous. Patient and clinician experiences with sharing data visualizations integrated into mental health treatment. *Social Sciences*, 2023.

[9] Victoria Clarke and Virginia Braun. Thematic analysis. *The Journal of Positive Psychology*, 2017.

[10] Kim Phuong Dao, Katrien de Cocker, Huong Ly Tong, A Baki Kocaballi, Clara Chow, and Liliana Laranjo. Smartphone-delivered ecological momentary interventions based on ecological momentary assessments to promote health behaviors: Systematic review and adapted checklist for reporting ecological momentary assessment and intervention studies. *JMIR mHealth and uHealth*, 2021.

[11] Evelien Dietvorst, Manon Hj Hilligers, Jeroen S Legerstee, Lianne P de Vries, Annabel Vreeker, and Loes Keijsers. Real-time personalized feedback in mhealth for adolescents. *DIGITAL HEALTH*, 2024.

[12] Mila Hall, Lias M. Lappenbusch, Emily Wiegmann, and Julian A. Rubel. Administration and policy in mental health and mental health services research. *Adm Policy Ment Health*, 2025.

[13] Xing He, Rui Zhang, Rubina Rizvi, Jake Vasilakes, Xi Yang, Yi Guo, Zhe He, Mattia Prosperi, Jinhai Huo, Jordan Alpert, and Jiang Bian. Aloha: developing an interactive graph-based visualization for dietary supplement knowledge graph through user-centered design. *BMC Medical Informatics and Decision Making*, 2019.

[14] Emma Holdsworth, Erica Bowen, Sarah Brown, and Douglas Howat. Client engagement in psychotherapeutic treatment and associations with client characteristics, therapist characteristics, and treatment factors. *Clinical Psychology Review*, 2014.

[15] Shagheyegh Azizian Kia, Lisette Wittkampf, Jacobine van Lankeren, and Pauline Janse. Motives of therapists for using routine outcome monitoring (rom) and how it is used by them in clinical practice: Two qualitative studies. *Administration and Policy in Mental Health and Mental Health Services Research*, 2025.

[16] Andrew A. McAleavey, Kim de Jong, Helene A. Nissen-Lie, James F. Boswell, Christian Moltu, and Wolfgang Lutz. Routine outcome monitoring and clinical feedback in psychotherapy: Recent advances and future directions. *Administration and Policy in Mental Health and Mental Health Services Research*, 2024.

[17] Marry L. McHugh. Interrater reliability: the kappa statistic. *Biochem Med*, 2012.

[18] M. R. Mehl and T. S. Conner. *Handbook of research methods for studying daily life*. The Guilford Press, 2012.

[19] Silvia Miksch and Wolfgang Aigner. A matter of time: Applying a data–users–tasks design triangle to visual analytics of time-oriented data. *Computers Graphics*, 2014.

[20] Inez Myin-Germeys, Zuzana Kasanova, Thomas Vaessen, Hugo Vachon, Olivia Kirtly, Wolfgang Viechtbauer, and Ulrich Reininghaus. The experience sampling methodology as a digital clinical tool for more person-centered mental health care: an implementation research agenda. *Psychological Medicine*, (1), 2024.

[21] Marilyn L. Piccirillo, Emorie D. Beck, and Thomas L. Rodebaugh. A clinician's primer for idiographic research: Considerations and recommendations. *Behavior Therapy*, 2019.

[22] Qualtrics. Qualtrics Survey Platform, 2024. Accessed: 2025-06-02.

[23] Jim van Os, Simone Verhangen, Anne Marsman, Frenk Peeters, Machteld Marcelis Maarten Bak, Marjan Drukker, Nele Jacobs Ulrich Reininghaus, Claudia Simons Tinke Lataster, ESM-MERGE Investigators PhD, Richel Lousberg, Peter C. Groot Sinan Gülöksüz, Carsten Leu AND, Wolfgang Vieghtbauer, and Philippe Delespaul. The experience sampling method as an mhealth tool to support self-monitoring, self-insight, and personalized health care in clinical practice: van os et al. *Depression and Anxiety*, (1), 2017.

[24] Jeroen Weermeijer, Glenn Kiekens, Martien Wampers, Peter Kuppens, and Inez Myin-Germeys. Practitioner perspectives on the use of the experience sampling software in counseling and clinical psychology. *Behaviour Information Technology*, 2024.

[25] Yuanyan Zhang, Hong Ma, Yubo Wang, Xiangju Du, and Danni Chi. Burnout and its associated factors in psychotherapists. *Current Psychology*, 2024.

[26] Ole Karkov Østergard, Lasse Grønnebæk, and Kristine Kahr Nilsson. Do therapists know when their clients deteriorate? an investigation of therapists' ability to estimate and predict client change during and after psychotherapy. *Clinical Psychology Psychotherapy*, 2024.