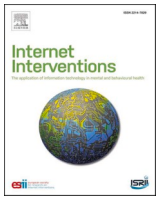
Internet Interventions 42 (2025) 100873



Contents lists available at ScienceDirect

Internet Interventions

journal homepage: www.elsevier.com/locate/invent



Evaluating visual feedback of experience sampling data for mental health practitioners

Maarten Piot [a,*], Egon Dejonckheere [a,b], Anke Tuinstra [a], Imke Tijs [a], Peter Kuppens [a], Stijn Verdonck [a]

[a] Faculty of Psychology and Educational Sciences, KU Leuven, Leuven, Belgium
[b] Department Medical and Clinical Psychology, Tilburg School of Social and Behavioral Sciences, Tilburg, Belgium

ARTICLE INFO

Keywords:
E-health
Clinical implementation
Experience sampling
Blended care
Personalized care
Visual feedback

ABSTRACT

Background: The Experience Sampling Method (ESM) allows to capture and quantify people's daily experiences, thoughts, and behaviors in real-world contexts. While ESM holds great promise for informing case conceptualization and improving therapeutic targeting in mental health care, visualizations of self-reported ESM data remain largely tailored to researchers and are often not easy to interpret for mental health practitioners. To what extent such visualizations can assist practitioners within their context of mental health care remains still unclear.
Objective: This study investigates the effect of including uncertainty information in visualizations of ESM feedback on practitioners' decision-making processes.
Methods: We recruited 40 Belgian mental health practitioners and students in their last year of psychology or educational science studies to take part in an experiment that assessed their proficiency in interpreting ESM feedback based on reality-inspired clinical cases. Three variables (i.e., interactive hover function, error bars, and textual descriptions about effect size and confidence interval) were manipulated in a balanced factorial design to determine their effects on correct data interpretation.
Results: The results indicate that providing textual descriptions about effect size and confidence intervals increased practitioners' confidence levels and accuracies compared to error bars alone. Furthermore, practitioners want to integrate ESM feedback into treatment. When ESM feedback was associated with low confidence, they preferred to further explore the clinical case. In contrast, when ESM feedback was associated with high confidence, they opted for more action-oriented interventions.
Conclusion: These findings underscore the importance of including textual descriptions in graphical visualizations as they help to improve practitioners' decision-making processes in mental health care.

1. Introduction

The Experience Sampling Method (ESM) is a research method that enables to capture and quantify people's thoughts, feelings, and behaviors in their daily lives (Bolger et al., 2003; Myin-Germeys and Kuppens, 2022). This method involves prompting individuals with questions multiple times a day on their smartphones, of which the responses can be used to inform therapeutic practice. By repeatedly measuring people's thoughts, feelings, and behavior, practitioners are given insight into patients' everyday patterns of emotions, behaviors, symptoms, context, and their interrelations over time (Myin-Germeys and Kuppens, 2022).

Aside from this more diagnostic use, ESM can also be used to evaluate the impact of interventions or for routine outcome monitoring (van Os et al., 2017). Feedback based on ESM data (hereafter referred to as 'ESM feedback') can thus provide many potential benefits for treatment. Leveraging ESM feedback empowers practitioners to adjust their treatment plans and select interventions (Hall et al., 2024; Kroeze et al., 2017). Collaborative discussions based on ESM feedback enhance patient empowerment and facilitate shared decision-making (Simons et al., 2015). For patients, receiving feedback can be motivational to continue monitoring (Hsieh et al., 2008), help them to deal with their symptoms (Folkersma et al., 2021), and introduce behavioral change (Snippe et al., 2016).

Realizing the potential of ESM hinges on its correct interpretation by practitioners. Recent advancements in online ESM platforms aim to assist practitioners in this task (Bos et al., 2022; Mestdagh et al., 2023;

* Corresponding author at: Faculty of Psychology and Educational Sciences, KU Leuven, Tiensestraat 102, Leuven, 3000, Belgium.
E-mail address: maarten.piot@kuleuven.be (M. Piot).

https://doi.org/10.1016/j.invent.2025.100873
Received 20 September 2024; Received in revised form 11 September 2025; Accepted 16 September 2025
Available online 17 September 2025
2214-7829/© 2025 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

Verhagen et al., 2017). Data visualizations play a crucial role for interpreting the ESM feedback, which can provide insights into patient trends and patterns. Weermeijer et al. (2023) suggest specific visualizations as clinically meaningful for practitioners, such as timelines showing symptom changes over time, bar charts displaying emotions in different contexts, and pie charts illustrating the frequency of specific activities (Weermeijer et al., 2023). According to our previous survey study with 89 practitioners and 64 researchers, visualizing variations in symptoms and experiences across contexts is especially useful (Piot et al., 2022).

Visualizations are important tools to communicate scientific data in a clear and engaging way (Midway, 2020). However, if visualizations are not well designed, they can be misleading for the readers to interpret and lead to biases (Mansoor and Harrison, 2018). Because interpretation can form the basis for treatment directions, it is of paramount importance that practitioners can interpret visualizations based on a correct assessment with minimal misinterpretation.

However, the interpretation of statistical test results remains challenging (Lytsy et al., 2022). To avoid misinterpretation, one of the requirements is incorporating uncertainty into graphical visualizations (e.g., line graphs, bar graphs) (Van Der Laan et al., 2015). Previous research has compared various methods for visualizing uncertainty in graphs, such as gradient plots, error bars or violin plots (Correll and Gleicher, 2014; Frans et al., 2023). In a recent study involving 230 practitioners, Frans et al. (2023) found that providing measurement uncertainty using traditional confidence intervals and error bars were perceived as most understandable, but some practitioners expressed a lack of visual support when only textual intervals were provided. This suggests that visualizing uncertainty can enhance the communication and comprehension of uncertainty for practitioners.

In this study, we examined the effect of including different types of (uncertainty) information in ESM feedback on practitioners' accuracy of their confidence assessment. We utilized fictitious, but real-life inspired, clinical cases of which practitioners were given visualizations presented in the m-Path dashboard (Mestdagh et al., 2023). m-Path is an online ESM platform developed collaboratively with practitioners and ESM researchers with the explicit aim to provide ESM feedback in the context of clinical practice (Dejonckheere et al., 2021). The dashboard featured a personalized feedback module presenting ESM feedback in dynamic graphical visualizations, allowing practitioners to interact with the data by moving their mouse over specific User Interface (UI) elements. This interaction, known as hovering, enables practitioners to access additional information without cluttering the interface. Hovering over data points in a timeline triggers a tooltip with contextual information, providing insight into fluctuations of stress levels, for instance, which can shed light on patient's social contact patterns. Similarly, hovering over peaks and valleys in a patient's health trajectory gives practitioners access to more relevant information without cluttering the interface.

Participants evaluated three graphical visualizations features: (1) interactive hovering function, (2) error bars, and (3) textual descriptions regarding effect size and confidence intervals. The graphical visualizations are illustrated in Fig. 1a. Participants were then questioned on interpretation correctness and had to rate the confidence of their interpretation of the ESM feedback. At the end of each clinical case, participants were asked about their treatment directions based on the presented cases in the dashboards. We addressed the following research questions:

1. How does the inclusion of uncertainty information in ESM feedback, such as error bars and textual descriptions regarding effect sizes and confidence intervals, affect practitioners' decision-making regarding statistical confidence?
2. How does the inclusion of uncertainty information in ESM feedback influence the accuracy of practitioners' confidence assessments?
3. How does practitioners' confidence in ESM feedback affect their decision-making regarding treatment directions?

This study will provide practical recommendations for visualizing uncertainty in ESM feedback to facilitate interpretations for mental health practitioners. Furthermore, we want to explore the challenges practitioners encounter when interpreting visualizations within a clinical context. As such, this study is intended to contribute to the development of ESM feedback in mental health care.

2. Methods

2.1. Participants

The study used convenience sampling to recruit Belgian mental health practitioners and students in their last year of psychology or educational sciences studies. Inclusion criteria were Dutch-speaking Master students in psychology or special education and mental health practitioners (e.g., psychologists, special needs educationalists). Non-Dutch speakers were excluded. Mental health practitioners were approached from a group practice in Leuven via newsletters and flyers

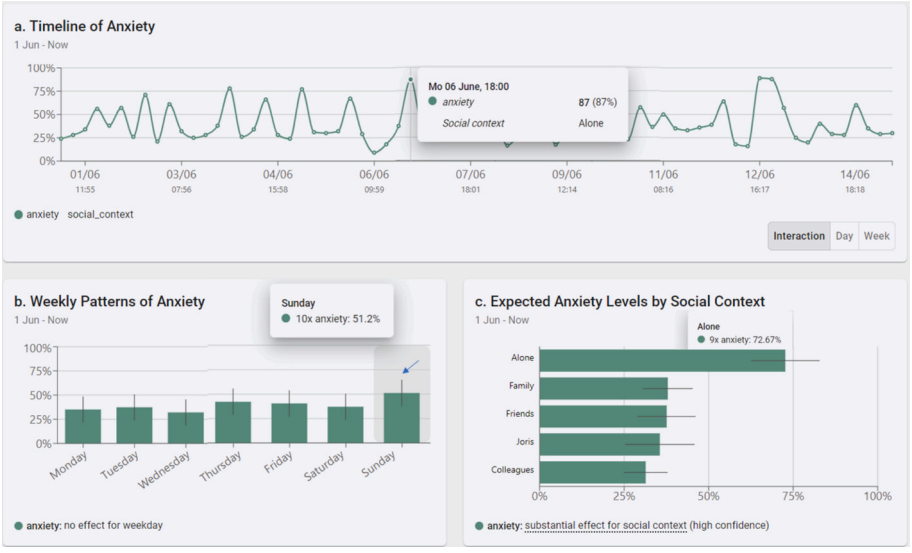


Fig. 1. Example of a personalized dashboard. This dashboard includes the following graphs: (a) timeline, (b) week patterns, and (c) expected levels. Visualization features are: interactive hover function, error bars, textual description about effect size and confidence intervals.

distributed during a workshop on using a mental health tool. Final-year psychology and educational sciences students were invited through online community groups. Neither practitioners nor students were required to have prior experience with or knowledge of ESM.

In total, 40 participants participated in the experiment between December 2021 and March 2022. Participants were primarily psychology Master's students, followed by mental health practitioners, including psychologists and special needs educationalists (see Table 1). In exchange for participating, participants who completed the experiment had the chance to win one of three Fnac vouchers. The study procedures were approved by the social and societal ethics committee of KU Leuven (no. G-2021-4504). We did not preregister our hypotheses. All participants signed an informed consent sheet.

2.2. Design

The study used a balanced design to investigate how visualization features affect the interpretation of statistical confidence in clinical practice. Participants interacted with computer dashboards that presented graphical visualizations of four reality-inspired clinical cases.

The cases included: (1) a child in special education with behavioral issues, (2) an elderly person experiencing loneliness in a service flat, (3) a student struggling with concentration issues, and (4) a child playing soccer as a coping mechanism for severe anger outbursts. The case examples were provided by IT and AT, both Master's students in psychology and educational sciences, respectively. The complete case descriptions can be found in Table A1 of Appendix A.

The software used in this study was built upon on the clinical m-Path dashboard (*M-Path*, n.d.) and further developed by MP, a Master's graduate in software engineering technology. The visualization features in the visual feedback were:

1. Hover Function: Additional information displayed upon hovering over a data point. This visualization feature allowed participants to explore more information about the ESM data. In the timeline, a tooltip provided additional contextual information and displayed exact scores of variables at specific time points, such as social contact and anxiety levels. For bar charts, a tooltip displayed exact frequencies of a variable for each category. If a variable had less than five measurements, a message with "too few measurements" was shown.
2. Error Bars: Graphical representation of 95 % confidence interval. Error bars were included at the top of each bar to indicate the boundaries of a 95 % confidence interval.
3. Textual Description: Textual description of effect size and statistical confidence. Below bar charts, a short message described the observed effect strength with corresponding confidence levels.

Table 1
Details participant characteristics.

Characteristics	$n = 40$
Gender, n [%]	
Male	8 [20]
Female	32 [80]
Age, years	
Mean (SD)	26.2 (7.3)
Profession, n [%]	
Psychologist	4 [10]
Special needs educationalist[a]	5 [12.5]
Master student educational sciences	8 [20]
Master student psychology	23 [57.5]
Experience with ESM[b] [mean (SD)]	3 (2.24)

Notes. Abbreviations: SD = standard deviation.
[a] The term "special needs educationalist" refers to a professional who specializes in special needs education and support for individuals with learning difficulties. In Dutch, this is known as "orthopedagoog".
[b] Rated from 1 ("No experience") to 7 ("Much experience").

Each visualization feature could be turned on (1) or off (0), resulting in 8 unique conditions: (0,0,0), (0,0,1), (0,1,0), (0,1,1), (1,0,0), (1,0,1), (1,1,0), (1,1,1).

2.3. Procedure

Participants were assigned to different conditions based on the combinations of the three visualization features. Each participant received all four clinical cases but with different combinations of the visualization features. Participants were divided into five blocks of eight, ensuring each condition was presented five times per case. The presentation order was counterbalanced using a balanced Latin square to reduce order effects (Bradley, 1958). An example assignment of the conditions in blocks 1 and 2 are presented in Table 2.

Before starting the experiment, participants watched a 5-minute instructional video, explaining the different graphs and visualization features. This instructional video is summarized in Fig. 1, which depicts the different graphs (e.g., timeline, week patterns, expected levels) and visualization features (i.e., interactive hover function, error bars, textual description). Participants then received instructions of the tasks, explaining statistical confidence as a measure indicating the likelihood of an effect exists when additional measurements are collected. To explain this further, the instructions of the task provided a practical example about anxiety levels as a function of social context (see Fig. 1c). If confidence is high, we can say with more confidence that the effect for social context is related to anxiety. In contrast, if confidence is low, it is possible that anxiety is not related to social context, but by chance or other factors.

To evaluate participants' proficiency in interpreting ESM feedback, each clinical case consists of a list of statements about the information that be gleaned in the dashboard visualizations. Participants were given a list of in total 13 statements (see Table A1 in Appendix A). These statements assessed the raw statistical differences between different measured variables and could be seen as an exercise in reading graphs. An example statement included: "Across all conducted measurements, Louise exhibits an on average higher anxiety score on Wednesdays compared to other weekdays". For each statement, participants were asked to:

1. Judge the veracity: Evaluate the correctness of the statement as true or false.
2. Rate the statistical confidence: Indicate the statistical confidence of the statement using a 5-point Likert scale ranging from 1 ("Very low confidence") to 5 ("Very high confidence").
3. Express personal certainty: Indicate their own certainty about their response using a 5-point Likert scale ranging from 1 ("Very uncertain") to 5 ("Very certain"). In the instructions of the task, certainty was defined as the extent to which participants felt assured about their response to a statement.

At the end of each clinical case, participants answered an open-ended

Table 2
Assignment of conditions in block 1 (left) and block 2 (right).

P nr.	Case 1	Case 2	Case 3	Case 4
1	111	000	100	011
2	000	011	111	101
3	011	101	000	001
4	101	001	011	010
5	001	010	101	110
6	010	110	001	100
7	110	100	010	111
8	100	111	110	000

P nr.	Case 1	Case 2	Case 3	Case 4
9	110	101	010	001
10	100	001	110	010
11	111	010	100	110
12	000	110	111	100
13	011	100	000	111
14	101	111	011	000
15	001	000	101	011
16	010	011	001	101

question about their conclusions and proposed treatment directions (i.e., "How would you use this visual feedback in treatment?").

To summarize, participants were given descriptions of four clinical cases along with ESM feedback. They were required to evaluate 13 statements about dashboard visualizations based on three characteristics: veracity, statistical confidence, and personal certainty.

The experiment was piloted with two Master's students and two ESM researchers to test the survey design, understand the questions, and estimate completion time, which averaged around 40 min. Participants were individually invited to a location in Leuven, where experimenter IT guided them through the task while experimenter AT recorded completion times and responses. Laptop screens and audio were recorded using the built-in Windows screen recorder. At the end of the experiment, participants provided demographic information (e.g., age, sex, profession), including their experience with mental health tools. Example questions included: "I have experience with digital tools for measuring treatment and/or making a diagnosis, such as apps or patient tracking systems." and "I have experience with the Experience Sampling Method." Responses were collected using a 7-point Likert scale ranging from 1 ("Strongly disagree") to 7 ("Strongly agree").

2.4. Data analysis

Data analysis was organized as a function of our three research questions:

RQ1: How does the inclusion of uncertainty in ESM feedback, such as error bars and textual descriptions regarding effect sizes and confidence intervals, affect practitioners' decision-making regarding statistical confidence and certainty?

We analyzed participants' responses on 5-point Likert scale statements regarding their statistical confidence and certainty. We used two-sample *t*-tests to examine whether enabling or disabling specific visualization features significantly affected confidence assessment and certainty, respectively. Additionally, a linear mixed-effects model (LMM) was applied to further analyze the effects of the visualization features on confidence assessment and certainty. The model included fixed effects for the hover function, error bars, and textual descriptions about effect size and confidence intervals. Random intercepts were included for participants and the cases they evaluated.

RQ2: How does the inclusion of uncertainty in ESM feedback influence the accuracy of participants' confidence assessment?

To establish a benchmark, two Master's students and two ESM researchers individually responded to the survey statements. After a collaborative discussion, we assigned an average consensus score to each statement. We then calculated the differences between participants' self-reported confidence scores and our consensus scores, with the absolute differences representing inaccuracies in participants' confidence assessment. An inaccuracy of zero representing a perfectly accurate confidence assessment (at least compared to the four raters' average consensus score). We used a two-sample *t*-test to test whether enabling or disabling specific visualization features significantly influenced the inaccuracy of confidence assessments. An LMM was also conducted to analyze the effects of the visualization features on the inaccuracy of the confidence assessments.

RQ3: How does practitioners' confidence in ESM feedback affect their decision-making regarding treatment directions?

Participants were asked an open-ended question regarding their treatment directions after each case. The responses were transcribed verbatim and analyzed through a qualitative content analysis in Excel, following an inductive coding approach (Savin-Baden and Howell Major, 2023). This method involved a bottom-up approach, where we did not use predefined categories but instead derived themes that emerged from the data. The main author (M.Sc., male, Ph.D. student) reviewed the responses for each participant and categorizing each one. These categories were then grouped and quantified into overarching themes.

For each participant, we calculated the mean confidence score in each clinical case, hereafter referred to "case confidence". Subsequently, we performed a Pearson correlation analysis to examine the relationship between the case confidence and the identified themes. Only themes with occurrences higher than 10 % were taken into account. An LMM was further conducted to analyze the effects of the identified themes on the case confidence.

3. Results

Participants evaluated the interpretation correctness of each statement on the raw statistical differences between different measured variables. They scored an average of 12 out of the total 13 statements correctly ($SD = 1.2$). In what follows, we present the results of each research question. First, we present the results of the confidence assessment, certainty, and inaccuracy scores for enabling and disabling each graphical feature. An overview of descriptive statistics for each graphical feature is shown in Table 3. Finally, we present the results of the treatment directions and compared them to case confidence.

3.1. RQ1: how does the inclusion of uncertainty in ESM feedback, such as error bars and textual descriptions regarding effect sizes and confidence intervals, affect practitioners' decision-making regarding statistical confidence and certainty?

3.1.1. Confidence assessment

When aggregating over participants and statements, confidence levels when enabling and disabling each visualization feature are shown in Fig. 2. When participants received textual descriptions, they indicated significantly higher confidence ($M = 2.87$) compared to the case without textual information ($M = 2.36$, $t = 4.37$, $p < .001$). This trend was consistent within participants, as confirmed by the mixed-effects model (see Table A3.1 of Appendix A). Notably, the presence of textual descriptions about effect size and confidence intervals was the only significant predictor of the confidence assessment ($\beta = 0.471$, $t = 4.20$, $p < .001$).

3.1.2. Certainty

When aggregating over participants and statements, certainty scores were significantly higher when participants had access to textual information ($M = 2.81$) compared to those without ($M = 2.50$, $t = 2.74$, $p = .0032$). Similarly, within participants, the mixed-effects model revealed that textual descriptions about effect size and confidence intervals was the only significant predictor of certainty ($\beta = 0.236$, $t = 2.12$, $p = .034$; see Table A3.2 of Appendix A).

3.2. RQ2: how does the inclusion of uncertainty in ESM feedback influence the accuracy of participants' confidence assessment?

When aggregating over participants and statements, inaccuracies for enabling and disabling each visualization feature is shown in Fig. 3. When participants had access to textual description about effect sizes and confidence intervals, they demonstrated significantly lower inaccuracies ($M = 1.04$) compared to no textual information ($M = 1.56$, $t = -4.86$, $p < .001$). The mixed-effects model supported this finding, indicating that textual descriptions was the only significant predictor of the inaccuracy of the confidence assessments ($\beta = -0.507$, $t = -4.81$, $p < .001$; see Table A3.3 of Appendix A).

3.3. RQ3: how does practitioners' confidence in ESM feedback affect their decision-making regarding treatment directions?

At the end of each case, participants were asked an open-ended

Table 3
Mean scores for enabling and disabling each graphical feature when aggregating over participants and statements.

	Hovering		Error bars		Textual information[a]	
	Enabled	Disabled	Enabled	Disabled	Enabled	Disabled
Confidence	2.62 (1.41)	2.61 (1.33)	2.68 (1.38)	2.55 (1.35)	2.87 (1.47)	2.36 (1.20)
Certainty	2.68 (1.32)	2.62 (1.30)	2.63 (1.32)	2.68 (1.30)	2.81 (1.38)	2.50 (1.21)
Inaccuracy	1.33 (1.30)	1.27 (1.17)	1.25 (1.19)	1.35 (1.28)	1.04 (1.21)	1.56 (1.21)

[a] All *p*'s for textual information were smaller than 0.05, calculated using a *t*-test and LMM.

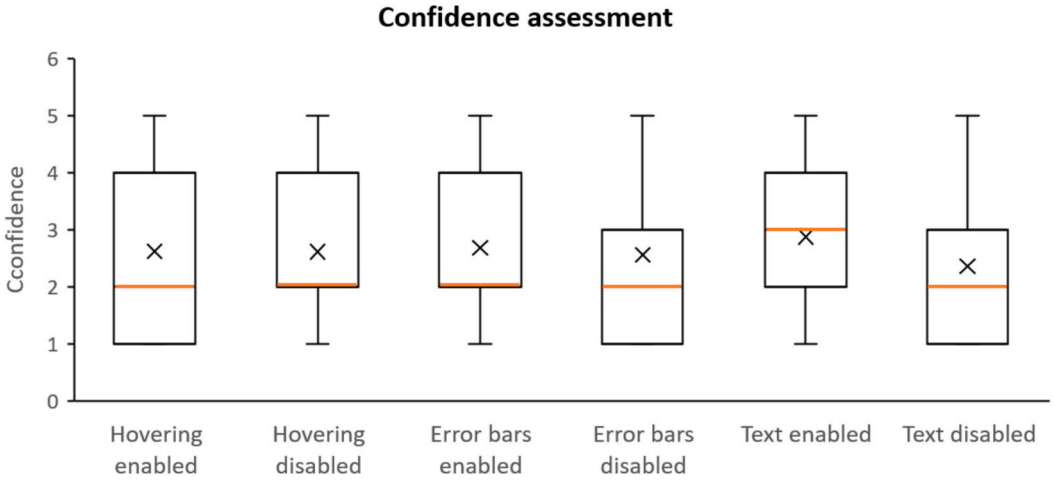


Fig. 2. Confidence levels for each visualization feature comparing enabled or disabled states.
Note. Orange horizontal lines represent the median.

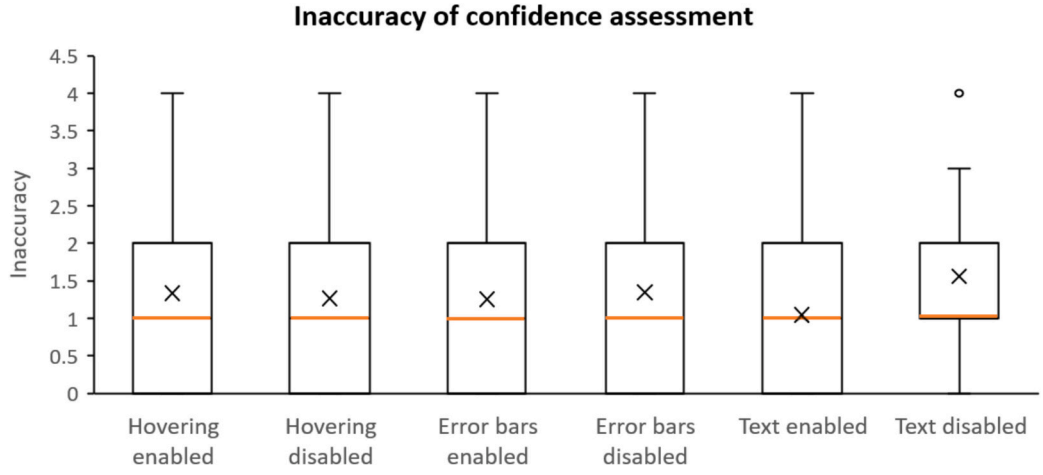


Fig. 3. Inaccuracy of the confidence assessment.
Note. The black horizontal line (inaccuracy = 0) indicates a perfectly accurate confidence assessment.

question about their treatment directions based on the presented ESM feedback (i.e., "How would you use this visual feedback in treatment?"). Almost all participants indicated they would use the provided ESM feedback to propose a further treatment direction ($n = 233$). Only one participant would not use the feedback. After coding participants' responses using the inductive approach, three overarching categories were identified, which were: further exploration ($n = 160$), action-oriented interventions ($n = 69$), and providing insights towards the patient or network ($n = 5$). Table 4 summarizes the results of each treatment direction using frequencies and percentage distribution. Example quotes are presented in Table A2 of Appendix A.

3.4. Further exploration

Participants expressed their willingness to explore the clinical case through quantitative or qualitative methodologies. The quantitative approach ($n = 49$) aimed to collect more measurement over a longer period by prolonging the questionnaire (quotes 1–7 in Table A2 of Appendix A). Some participants suggested creating a new personalized questionnaire to investigate other factors (quotes 8 and 9 in Table A2 of Appendix A).

Qualitative methods ($n = 102$) aimed to provide a better understanding of the case, as the available evidence was deemed insufficient for participants to formulate conclusive assessments. Qualitative exploration involved conducting interviews with patients ($n = 43$,

Table 4
Frequencies and percentage distribution of treatment directions.

Treatment direction	Frequency (*n*)	Percentage (%)
Further exploration	160	68.4
➢ Quantitative exploration	49	20.9
➢ Qualitative exploration	102	43.6
◦ Patient interviews	63	26.9
◦ Exploring patient's network	28	12.0
◦ Observation	11	4.7
➢ Exploration (not specified)	9	3.8
Action-oriented interventions	69	29.5
➢ Interventions involving patient's network	28	11.2
➢ Interventions in physical activities	23	9.8
➢ Interventions targeting psychological issues	18	7.7
Providing insights towards the patient or network	5	2.1
No treatment direction proposed	1	0.4

quotes 10–15 in Table A2 of Appendix A) or their network ($n = 28$, quotes 15 and 16 in Table A2 of Appendix A), and collecting observational data ($n = 11$, quote 17 in Table A2 of Appendix A).

3.5. Action-oriented interventions

Participants proposed action-oriented interventions ($n = 69$) where patients took concrete action in self-managing their well-being. Possible interventions included: involving patient's network such as family and friends into the treatment process ($n = 28$, quotes 18–21 in Table A2 of Appendix A), engaging in physical activities ($n = 23$, quotes 22–24 in Table A2 of Appendix A) and targeting psychological issues ($n = 18$, quotes 25–27 in Table A2 of Appendix A).

3.6. Providing insight towards the patient or network

Participants aimed to provide the patient or their network with insights by identifying factors that promote self-reliance, thereby reducing reliance on specific coping mechanisms ($n = 5$, quotes 28–31 in Table A2 of Appendix A).

3.7. Comparing treatment directions to case confidence

Treatment directions regarding further exploration (68.4 %), whether qualitative or quantitative, and action-oriented interventions (29.5 %) were compared towards case confidence. When aggregating over participants and cases, we observed a negative correlation between case confidence and a desire for further exploration of the clinical case ($r = -0.20$, $t = -2.53$, $p = .012$). This suggests that as their confidence increased, they were less likely to recommend further exploration. Similarly, we observed a positive correlation between case confidence and a desire for taking action based on the clinical case ($r = 0.23$, $t = 2.99$, $p = .0033$). Higher confidence was associated with greater likelihood of recommending action-oriented interventions. The mixed-effects model confirmed this result, revealing a negative effect of case confidence on the likelihood of recommending further exploration ($\beta = -0.126$, $t = -2.35$, $p = .020$; see Table A3.4 of Appendix A) and a positive effect of case confidence on the likelihood of recommending action-oriented interventions ($\beta = 0.126$, $t = 2.63$, $p = .009$; see Table A3.5 of Appendix A).

4. Discussion

In this study, we evaluated whether practitioners could make informed decisions based on the visualizations provided by the ESM dashboard. By incorporating uncertainty such as error bars, and textual descriptions about effect sizes and confidence intervals, we aimed to elucidate the influences of these features on practitioners' confidence assessment and accuracy. Participants were able to correctly interpret the raw statistical differences between different measured variables. The inclusion of textual descriptions significantly increased their confidence assessment. In particular, when participants had access to textual descriptions, they reported higher confidence compared to those without. Similarly, when participants are provided with textual descriptions, they showed significantly higher accuracies in their confidence assessment compared to those without.

However, the presence of error bars did not influence participants' statistical confidence, nor accuracy. This suggests that error bars posed greater challenges for participants in drawing accurate conclusions regarding statistical confidence in clinical cases. This finding contrasts with a previous study conducted by Frans et al. (2023), where practitioners successfully interpreted error bars to assess probabilities. Our study diverged from previous research in that our participants were not provided with any cutoff points for assessing their statistical proficiency through mathematical exercises. Instead, we presented participants with clinical case scenarios, formalizing the clinical context within an experiment that aimed to capture not just their ability to correctly solve confidence assessments, but also to incorporate clinical context in the decision-making processes. When it comes to making decisions about confidence, relying solely on error bars for interpreting uncertainty may prove insufficient. Professionals who are more familiar with statistics may favor exact representations of visualizations communicating uncertainty, while less familiar professionals prefer visualizations that convey uncertainty in a more intuitive or semantic matter (Franconeri et al., 2021). Consequently, textual descriptions align more closely with practitioners' perceptions compared to traditional error bars. As such, effectively communicating uncertainty in mental health care may require using textual descriptions to guide practitioners towards accurate conclusions regarding statistical confidence.

Our findings revealed that providing ready-to-use ESM feedback equipped practitioners with the ability to directly interpret ESM data, make informed decisions, and take further steps based on ESM feedback. In line with findings from prior research (Bos et al., 2019; Kroeze et al., 2017; Weermeijer et al., 2023), practitioners acknowledged the underlying benefits of integrating ESM feedback into treatment because they wanted to act based on the provided ESM feedback in our clinical cases. Practitioners endorsed further exploration (68.4 %), and action-oriented interventions (29.5 %) as the preferred treatment directions. Interestingly, the least frequently mentioned treatment direction was providing insights to the patient or their network (2.1 %). This finding somewhat contrasts with the responses of our earlier quantitative survey, where practitioners endorsed giving patients access to ESM feedback (Piot et al., 2022). In the present mixed-method study, however, this theme emerged inductively during the coding process rather than being asked directly in a questionnaire. In line with another mixed-method study, Weermeijer et al. (2023) found that while practitioners generally value the use of ESM in clinical practice, they also recognize that its applicability is not universal. Similarly, in the qualitative study by Bos et al. (2019), interviewed clinicians believed that most patients could benefit from ESM, but they did not consider it equally useful for all patient populations. Taken together, this body of research suggests that the usefulness of ESM feedback is not a one-size-fits-all solution, but should be evaluated case by case, depending on the clinical context and patient symptomatology.

Our findings align closely with Hill (2019) three-stage model of helping, providing a holistic perspective on the therapeutic journey through three stages: (1) exploration, (2) insight, and (3) action. Firstly, participants wanted to further explore the ESM data, reflecting the client-centered approach in Hill's exploration stage. This involves delving deeper into patients' experiences through qualitative methodologies, such as patient interviews and observations. Secondly, the insight and awareness stage built upon the exploration stage by providing patients with insights. Participants aimed to give patients a deeper understanding of underlying issues and promote self-reliance through these insights. Thirdly, participants emphasized the

importance of action-oriented interventions, which correspond to Hill's action stage. These interventions incorporate Cognitive and Behavior theories into treatment strategies, such as leveraging the patient's support network (e.g., establishing safety nets, enhancing familial bonds), engaging in physical activities (e.g., playing sport), and addressing psychological issues through skill development (e.g., aggression regulation). Overall, ESM feedback can assist practitioners across all three stages of the helping model.

When it comes to designing graphical visualizations, including uncertainty in graphs plays a crucial role in decision-making (Van Der Laan et al., 2015). Whether induced through textual descriptions or error bars, uncertainty prompts practitioners to either make action-oriented interventions or delve deeper into the ESM data for additional exploration. Higher confidence scores reflected participants to choose more action-oriented interventions, while participants with lower confidence scores preferred additional exploration. Because uncertainty influences clinical decision-making, it is recommended to integrate confidence information in textual form within ESM feedback. This approach can improve practitioners' clinical decision-making processes, with potential implications on the speed of moving between data exploration to proposing interventions in treatment.

4.1. Limitations

Several limitations should be acknowledged when interpreting the study's results. First, the absence of patient involvement in interpreting the ESM feedback may limit the generalizability of the findings, as feedback to both patients and practitioners could yield greater efficacy compared to feedback to practitioners alone (De Jong et al., 2014). Therefore, future research should investigate the spontaneous and actual use of ESM in mental health care. This includes exploring how patients and practitioners engage with ESM feedback in real clinical contexts and whether such use impacts treatment decisions. Second, most of our participants received statistical training as part of their psychology Master's degree, potentially skewing the results towards participants with higher statistical proficiency. Adjustments to textual descriptions may be necessary for participants with less statistical training. Third, only a short textual description was displayed in the graphical visualizations, but the inclusion of additional information can quickly become too complex, leading to information overload (Merlo and Hawamdeh, 2022). Finally, while expressing uncertainty in graphical visualizations is important, practitioners must also accurately interpret conclusions drawn from ESM feedback, recognizing that statistic significant results do not necessary implies clinical significance (Page, 2014).

5. Conclusions

The study sheds light on practitioners' interpretation skills regarding graphical visualization features in ESM feedback and their influence on treatment directions in mental health care. When participants had access to textual descriptions about effect sizes and confidence intervals, they demonstrated improved accuracies and higher confidence scores compared to those without. Error bars presented challenges for practitioners in interpretation, highlighting the necessity of directly integrating uncertainty into accompanying text for effective communication with practitioners. Furthermore, almost all participants indicated they would use the provided ESM feedback in treatment. Participants indicating higher confidence scores were less inclined to further exploration of the clinical cases and mentioned more action-oriented interventions, emphasizing the importance of incorporating uncertainty in ESM feedback to enhance decision-making processes. Future research should investigate the actual, spontaneous use of ESM feedback on treatment outcomes.

Abbreviation

ESM Experience sampling method

CRediT authorship contribution statement

Maarten Piot: Writing – review & editing, Writing – original draft, Supervision, Project administration, Visualizations, Software, Methodology, Investigation, Formal analysis, Conceptualization.
Egon Dejonckheere: Writing – review & editing, Supervision, Conceptualization.
Anke Tuinstra: Writing – review & editing, Investigation, Data collection, Clinical cases.
Imke Tijs: Writing – review & editing, Investigation, Data collection, Clinical cases.
Peter Kuppens: Writing – review & editing.
Stijn Verdonck: Writing – review & editing, Supervision, Software, Methodology, Investigation, Conceptualization.

Declaration of Generative AI and AI-assisted technologies in the writing process

During the preparation of this work the main author used ChatGPT in order to improve the readability and language of the manuscript. After using this tool, the authors reviewed and edited the content as needed and take full responsibility for the content of the published article.

Funding

This study was financially supported by the “KU Leuven technology transfer grant” (C3/3H200318; Dr. Kuppens). Dr. Dejonckheere is a post-doctoral research fellow supported by the Research Foundation Flanders (FWO; grant no. 1210621N). The funding agencies had no role in the interpretation of the results and the drafting of this manuscript.

Declaration of competing interest

The authors declare the following financial interests and personal relationships which may be considered as potential conflicts of interest: ED and SV are co-founders of m-Path, the commercial experience sampling software that was used in this study. PK is shareholder and member of the m-Path board of directors. MP is shareholder and will join the company as a user experience researcher.

Acknowledgements

We thank all participants for their valuable insights.

Appendix A

References

Bolger, N., Davis, A., Rafaeli, E., 2003. Diary methods: capturing life as it is lived. Annu. Rev. Psychol. 54, 579–616. https://doi.org/10.1146/ANNUREV.PSYCH.54.101601.145030.

Bos, F.M., Snippe, E., Bruggeman, R., Wichers, M., van der Krieke, L., 2019. Insights of patients and clinicians on the promise of the experience sampling method for psychiatric care. Psychiatr. Serv. 70 (11), 983–991. https://doi.org/10.1176/APPI.PS.201900050.

Bos, F.M., von Klipstein, L., Emerencia, A.C., Veermans, E., Verhage, T., Snippe, E., Doornbos, B., Hadders-Prins, G., Wichers, M., Riese, H., 2022. A web-based application for personalized ecological momentary assessment in psychiatric care: user-centered development of the PETRA application. JMIR Ment. Health 9 (8), E36430. https://doi.org/10.2196/36430.

Bradley, J.V., 1958. Complete counterbalancing of immediate sequential effects in a Latin square design. J. Am. Stat. Assoc. 53 (282), 525. https://doi.org/10.2307/2281872.

Correll, M., Gleicher, M., 2014. Error bars considered harmful: exploring alternate encodings for mean and error. IEEE Trans. Vis. Comput. Graph. 20 (12), 2142–2151. https://doi.org/10.1109/TVCG.2014.2346298.

De Jong, K., Timman, R., Hakkaart-Van Roijen, L., Vermeulen, P., Kooiman, K., Passchier, J., Van Busschbach, J., 2014. The effect of outcome monitoring feedback to clinicians and patients in short and long-term psychotherapy: a randomized controlled trial. Psychother. Res. 24 (6), 629–639. https://doi.org/10.1080/10503307.2013.871079.

Dejonckheere, E., Verdonck, S., Piot, M., Mestdagh, M., 2021. KU Leuven lanceert m-Path: Het digitale verlengstuk voor uw dagelijkse behandelingen. Medisfeer 692, 21–24. https://research.tilburguniversity.edu/en/publications/ku-leuven-lanceert-m-path-het-digitale-verlengstuk-voor-uw-dageli.

Folkersma, W., Veerman, V., Ornée, D.A., Oldehinkel, A.J., Alma, M.A., Bastiaansen, J. A., 2021. Patients' experience of an ecological momentary intervention involving self-monitoring and personalized feedback for depression. Internet Interv. 26, 100436. https://doi.org/10.1016/J.INVENT.2021.100436.

Franconeri, S.L., Padilla, L.M., Shah, P., Zacks, J.M., Hullman, J., 2021. The science of visual data communication: what works. Psychol. Sci. Public Interest 22 (3), 110–161. https://doi.org/10.1177/15291006211051956. /ASSET/IMAGES/LARGE/10.1177_15291006211051956-FIG17.JPEG.

Frans, N., Hummelen, B., Albers, C.J., Paap, M.C.S., 2023. Visualizing uncertainty to promote clinicians' understanding of measurement error. Assessment 30 (8), 2449. https://doi.org/10.1177/10731911221147042.

Hall, M., Lappenbusch, L.M., Wiegmann, E., Rubel, J.A., 2024. To use or not to use: exploring therapists' experiences with pre-treatment EMA-based personalized feedback in the TheraNet Project. Admin. Pol. Ment. Health. https://doi.org/10.1007/S10488-023-01333-3.

Hill, C.E., 2019. Helping skills: facilitating exploration, insight, and action. In: Helping Skills: Facilitating Exploration, Insight, and Action, 5th ed. https://doi.org/10.1037/0000147-000

Hsieh, G., Li, I., Dey, A., Forlizzi, J., Hudson, S.E., 2008. Using visualizations to increase compliance in experience sampling. In: UbiComp 2008 - Proceedings of the 10th International Conference on Ubiquitous Computing, pp. 164–167. https://doi.org/10.1145/1409635.1409657.

m-Path. Retrieved November 24, 2021, from. https://m-path.io/landing/ (n.d.).

Kroeze, R., van der Veen, D.C., Servaas, M.N., Bastiaansen, J.A., Voshaar, R.C.O.V., Borsboom, D., Ruhe, H.G., Schoevers, R.A., Riese, H., 2017. Personalized feedback on symptom dynamics of psychopathology: a proof-of-principle study. J. Person Oriented Res. 3 (1), 1–11. https://doi.org/10.17505/JPOR.2017.01.

Lytsy, P., Hartman, M., Pingel, R., 2022. Misinterpretations of P-values and statistical tests persists among researchers and professionals working with statistics and epidemiology. Ups. J. Med. Sci. 127, 8760. https://doi.org/10.48101/UJMS.V127.8760.

Mansoor, H., Harrison, L., 2018. Data visualization literacy and visualization biases: cases for merging parallel threads. In: Cognitive Biases in Visualizations, pp. 87–96. https://doi.org/10.1007/978-3-319-95831-6_7/FIGURES/3.

Merlo, T.R., Hawamdeh, S., 2022. Information overload and the use of data analytics and visualization tools in organizations. In: Understanding, Implementing, and Evaluating Knowledge Management in Business Settings, pp. 168–188. https://doi.org/10.4018/978-1-6684-4431-3.CH009.

Mestdagh, M., Verdonck, S., Piot, M., Niemeijer, K., Kilani, G., Tuerlinckx, F., Kuppens, P., Dejonckheere, E., 2023. m-Path: an easy-to-use and highly tailorable platform for ecological momentary assessment and intervention in behavioral research and clinical practice. Front. Dig. Health 5, 1182175. https://doi.org/10.3389/FDGTH.2023.1182175/BIBTEX.

Midway, S.R., 2020. Principles of effective data visualization. Patterns 1 (9), 100141. https://doi.org/10.1016/J.PATTER.2020.100141.

Myin-Germeys, I., Kuppens, P. (Eds.), 2022. The Open Handbook of Experience Sampling Methodology: A Step-by-step Guide to Designing, Conducting, and Analyzing ESM Studies, 2nd ed. Center for Research on Experience Sampling and Ambulatory Methods Leuven, Leuven.

Page, P., 2014. Beyond statistical significance: clinical interpretation of rehabilitation research literature. Int. J. Sports Phys. Ther. 9 (5), 726. /pmc/articles/PMC4197528/.

Piot, M., Mestdagh, M., Riese, H., Weermeijer, J., Brouwer, J.M.A., Kuppens, P., Dejonckheere, E., Bos, F.M., 2022. Practitioner and researcher perspectives on the utility of ecological momentary assessment in mental health care: a survey study. Internet Interv. 30, 100575. https://doi.org/10.1016/J.INVENT.2022.100575.

Savin-Baden, M., Howell Major, C., 2023. Qualitative research: the essential guide to theory and practice. Qual. Res. https://doi.org/10.4324/9781003377986.

Simons, C.J.P., Hartmann, J.A., Kramer, I., Menne-Lothmann, C., Höhn, P., van Bemmel, A.L., Myin-Germeys, I., Delespaul, P., van Os, J., Wichers, M., 2015. Effects of momentary self-monitoring on empowerment in a randomized controlled trial in patients with depression. Eur. Psychiatry 30 (8), 900–906. https://doi.org/10.1016/J.EURPSY.2015.09.004.

Snippe, E., Hartmann, J.A., Kramer, I., Viechtbauer, W., Delespaul, P., Myin-Germeys, I., Simons, C.J.P., Menne-Lothmann, C., Booij, S.H., Wichers, M., 2016. Change in daily life behaviors and depression: within-person and between-person associations. Health Psychol. 35 (5), 433–441. https://doi.org/10.1037/HEA0000312.

Van Der Laan, D.J., De Jonge, E., Solcer, J., 2015. Effect of displaying uncertainty in line and bar charts: presentation and interpretation. In: IVAPP 2015 - 6th International Conference on Information Visualization Theory and Applications; VISIGRAPP, Proceedings, pp. 225–232. https://doi.org/10.5220/0005300702250232.

van Os, J., Verhagen, S., Marsman, A., Peeters, F., Bak, M., Marcelis, M., Drukker, M., Reininghaus, U., Jacobs, N., Lataster, T., Simons, C., Lousberg, R., Gülöksüz, S., Leue, C., Groot, P.C., Viechtbauer, W., Delespaul, P., 2017. The experience sampling method as an mHealth tool to support self-monitoring, self-insight, and personalized health care in clinical practice. Depress. Anxiety 34 (6), 481–493. https://doi.org/10.1002/DA.22647.

Verhagen, S.J.W., Berben, J.A., Leue, C., Marsman, A., Delespaul, P.A.E.G., Van Os, J., Lousberg, R., 2017. Demonstrating the reliability of transdiagnostic mHealth Routine Outcome Monitoring in mental health services using experience sampling technology. PLoS One 12 (10), e0186294. https://doi.org/10.1371/JOURNAL.PONE.0186294.

Weermeijer, J.D.M., Wampers, M., de Thurah, L., Bonnier, R., Piot, M., Kuppens, P., Myin-Germeys, I., Kiekens, G., 2023. Usability of the experience sampling method in specialized mental health care: pilot evaluation study. JMIR Form. Res. 7 (1), e48821. https://doi.org/10.2196/48821.