communications psychology

A Nature Portfolio journal

Perspective

https://doi.org/10.1038/s44271-026-00416-9

Experience sampling methods require more than numbers

Check for updates

Laura F. Bringmann[1] ✉, Guðrún R. Guðmundsdóttir[2], Leonie Schorrlepp[3,4], Leonie V. D. E. Vogelsmeier[4,5], Melissa M. De Smet[5,6], Charlotte Vrijen[7], Eeske van Roekel[3,4], Jessica Fritz[8], Jennifer X. Chen[5], Olivia Kirtley[9], Markus I. Eronen[10], Joanne M. Chung[11], Marie Stadel[1,4,5,12] & Katie Hoemann[13]

The experience sampling method (ESM) collects people's real-time reports about their feelings, actions, and surroundings. While this method originally included both numerical and open-ended responses, most studies today focus only on the numbers. We argue that ESM researchers should collect and analyze open-ended responses again, as they are crucial for understanding what the numbers really mean, and for capturing parts of experience that numbers alone cannot, such as context, the "why" behind responses, and the temporal order of events. Open-ended responses can improve ESM data by grounding it in real-world experiences and phenomena as they are experienced in everyday life. Therefore, handbooks and guidelines on ESM should again include dedicated sections on collecting and analyzing open-ended text items. We highlight future work that is needed to achieve the systematic integration of open-ended items and their analysis into ESM research.

The experience sampling method (ESM)[1], also known as ecological momentary assessment (EMA)[2,3] or ambulatory assessment (AA)[4], is a research approach designed to capture individuals' experiences, behaviors, and contexts repeatedly in their everyday environment. Participants typically complete short questionnaires daily or several times per day (usually on smartphones) over an extended period of time (often weeks or months). The use of ESM is particularly prevalent in disciplines such as psychology, health sciences, organizational behavior, and human-computer interaction, where understanding individuals' context-specific experiences in everyday life is crucial. This methodology has become an influential and rapidly growing area of study[5]. The type of data gathered typically includes individuals' self-reported experiences of emotions, cognitions, behaviors, physiological sensations, and features of their environmental context (e.g., where they occurred and with whom).

From its inception, ESM was designed to collect both quantitative data and open-ended responses, allowing participants to share their experiences, behaviors, social contexts, and thoughts in their own words[1]. For example: "What caused you the most anger today? Briefly describe"[6]. Or: "What have you been doing for the past 15 min?"[7]. Responses to these kinds of *open-ended items* (see Box 1 for a definition and examples, and Table 1 for an overview of open-ended ESM items) result in what we refer to as *qualitative data*. Over time, however, this kind of qualitative data has gradually been set aside in the field, which has increasingly focused on collecting responses to *closed-ended items* that are numeric or easily translated into numbers, including dichotomous scales, Likert scales, visual analog scales, and predefined categories[8,9] — what we refer to as *quantitative data*[10]. Initially, this was partly due to technical reasons, as it was difficult to fill in open-ended items with the personal digital assistants (PDAs) that were used for data collection in the 1990s and 2000s[11]. This shift to quantitative data is also clearly reflected in ESM handbooks and overview papers. While earlier handbooks and review articles (e.g., Hektner et al.[11] or Bolger et al.[12]) extensively cover the collection and coding of open-ended responses, more recent overviews and handbooks (e.g., Mehl & Conner[8], Bolger & Laurenceau [13], or Trull & Ebner-Priemer[4,14,15]) provide minimal substantive discussion, reflecting the narrowing focus on quantitative information.

[1]Department of Psychometrics and Statistics, University of Groningen, Groningen, The Netherlands. [2]Department of Clinical Psychological Science, Maastricht University, Maastricht, The Netherlands. [3]Department of Developmental Psychology, Tilburg University, Tilburg, The Netherlands. [4]Tilburg Experience Sampling Center, Tilburg University, Tilburg, The Netherlands. [5]Department of Methodology and Statistics, Tilburg University, Tilburg, The Netherlands. [6]Department of Methods of Educational Sciences, University of Leuven, Leuven, Belgium. [7]Department of Developmental Psychology, University of Groningen, Groningen, The Netherlands. [8]Department of Psychology, University of Osnabrück, Osnabrück, Germany. [9]Department of Psychology, University of Leuven, Leuven, Belgium. [10]Department of Theoretical Philosophy, University of Groningen, Groningen, The Netherlands. [11]Department of Psychology, University of Toronto Mississauga, Mississauga, Canada. [12]Department of Sociology, University of Groningen, Groningen, The Netherlands. [13]Department of Psychology, University of Kansas, Lawrence, KS, USA. ✉e-mail: l.f.bringmann@rug.nl

Box 1 | Open-ended ESM items

Definition	ESM items that provide open space for freely typed text, or for recording audio that is later transcribed into text.	
Examples	Text entry	Audio entry
		
fully open standalone	“What have you done since the previous notification?”[89]	“Describe where you are, what you are doing, and who you are with”[27]
fully open follow-up	“What was going on, just before you thought about hurting yourself?”[23]	“What is the best or most enjoyable part of this experience?”[27]
Open-ended in disguise	ESM items that are open-ended in form, but are actually used to gather quantitative data, such as the *number of times* a type of behavior was performed on that day. Such items are not the focus of this article.	
Items that involve video recordings, photographs, or drawings could also be seen as open-ended items. However, such approaches remain uncommon and are not the focus of this article.		

Despite this apparent decline in methodological attention, open-ended responses continue to be collected in many ESM studies. Our targeted review of published ESM studies and a scan of the ESM item repository (https://esmitemrepositoryinfo.com/; see the Supplementary Note for further details: https://osf.io/jemrp/files/tv94p) indicate that studies have included a broad variety of open-ended items, ranging from more general questions about momentary feelings or recent events to more targeted questions about why participants missed an assessment (i.e., beep or ping[16]; see Table 1 for examples). Many studies prompt participants to recall the most important or emotionally salient event since the last beep[17], or to anticipate an upcoming event[18,19]. Some studies have focused on specific kinds of events, such as identity-related experiences[20] or daily emotional triggers[6]. Open-ended items are also used to capture immediate context, such as what participants are doing, where they are, and who they are with, as well as concurrent thoughts[21] and feelings[22]. In some cases, prompts are used to query particular phenomena, such as self-harm[23] or travel experiences[24]. Relatedly, open-ended items are sometimes used to probe response processes and other measurement-related matters, such as asking participants to explain their ratings[25,26].

The most typical format for open-ended items is free-text boxes[17], although some invite short phrases or lists[22]. Open-ended items can be stand-alone[21], or follow up on participants’ responses to closed-ended items[25]. Responses to open-ended items are most often typed, but several studies have collected voice recordings[24,25,27] or combined open-text data with images, such as photos of meals[28,29] (Box 1). Some studies have also used adaptive, self-learning items that participants can personalize by adding their own response categories[30,31]. For example, in response to an item about activities, a participant might add options such as “hanging out with my best friend”, “soccer practice”, or “having patient consultations”, which they can then simply select in future prompts when applicable.

However, open-ended items are seldom analyzed. In our targeted review, we found over 150 ESM publications that included open-ended items, but only 40 that also analyzed and reported them, with 18 of these published after 2020 (Supplementary Note). While this may reflect a renewed interest in open-ended items, the widespread practice of collecting but not analyzing qualitative data warrants reflection. Such practice implies suboptimal resource use[32] and even raises ethical questions. Open-ended items increase participant burden, which may affect compliance and data quality[33] (though see refs. 34 and [35]). Such a burden may be justified if the data contribute meaningfully to scientific understanding, but it is difficult to justify when open-ended responses are left unanalyzed. At the same time, open-ended responses can lead to valuable insights, and failing to examine them risks losing information that could enrich our understanding of daily life experiences.

We adopt a history and philosophy of science perspective to examine the perks and pitfalls of quantitative methods and to consider the potential of qualitative data to provide detailed insight into experiences in daily life. As there are very few resources explicitly addressing the benefits of open-ended items, we present a case for the use of open-ended responses in ESM. We propose that the recent reemergence of interest in using open-ended ESM items should herald a revival of the method, and discuss what is needed to achieve the systematic integration of open-ended items and their analysis into ESM research.

The enduring influence of quantification

The current literature about ESM has a clear quantitative focus. Review papers, handbooks, and overview articles focus on quantitative data, often collected through Likert or visual analog scales, and on statistical models used to analyze these data[8,9,13,36]. For example, valence can be repeatedly measured over time (e.g., with 0–100 visual analog scales) in order to get insight into affect dynamics, which are then typically modeled using multilevel approaches and (multilevel) vector autoregressive techniques[13,37–40]. These models are then used to examine how features of these dynamics (e.g., autocorrelation or inertia) relate to outcomes such as depression[41,42]. Even suggestions on how to improve ESM measures focus on visual analog or Likert scales and statistics[5,16,43–45]. As a result, both reliability and validity are treated primarily as statistical problems: reliability is assessed through increasingly sophisticated statistical models[46–49], and validity is typically

Table 1 | Overview of current practices using open-ended items in ESM

Study element	Decision point	Representative options	Description	Example publication(s)
Data collection	What information to capture?	General event descriptions	Participants recall the most important or emotionally salient event since the last assessment	17,120
		Specific types of events	Participants only report on particular kinds of experiences	20,23
		Immediate context and experience	Participants describe what they are doing, where they are, who they are with, current thoughts and feelings	21
		Response processes	Participants explain how or why they responded as they did	25
	What format to use?	Fully open	Participants are prompted for freely-generated, running text	17
		Fully open (list)	Participants are prompted to provide short phrases or lists	22
		Partially open	Participants are provided with multiple-choice items, including an editable 'other' option	121
	How to integrate with other items?	Standalone	Responses can be interpreted independently of other items	21
		Follow-up	Responses expand upon responses to other, closed-ended items	25
	What type of data to collect?	Typed text	Participants write their responses using a computer or phone keyboard	17
		Voice recordings	Participants speak their responses into a phone or recording device	24
		Non-linguistic data	Participants upload additional media (e.g., images)	28
Data analysis	How will data be processed?	Deductive manual coding	Researchers classify participant responses based on established frameworks or prior work	17
		Inductive manual coding	Researchers derive response categories from the data	25
		Combination of deductive and inductive	Researchers add new response categories, based on the data, to a predefined list	6
		Automated coding	Responses are classified or scored using Natural Language Processing techniques, such as word counting programs and Large Language Models	89,120
	How will data be analyzed?	Descriptive statistics	Raw or coded responses are quantitatively characterized	27,122
		Inferential statistics	Coded responses are used to predict outcomes of interest or compare across conditions/groups	6,112
		Qualitative interpretation	Researcher observations are verbally described, guided by example quotes	123
	How will data be presented?	Illustrative quotes	Raw responses are selected to highlight key observations	123
		Visualizations	Distributions of coded responses are illustrated using pie charts, word clouds, Venn diagrams, etc.	22

examined through statistical associations with other quantitative measures[50,51].

The quantitative focus in ESM research aligns with the general tradition in psychology, where, already in the late 19th century, leading figures such as Gustav Fechner and Wilhelm Wundt were adamant that psychological phenomena could and should be quantified[52]. This belief was influenced by the long history of successes that quantification had brought to the physical sciences in previous centuries. In turn, this led to the idea that quantitative science is real science, making it even more important for psychologists to quantify their subject matter in order to be recognized as a science[52–54]. This tradition was particularly strong in early psychometrics, where the focus was on developing standardized tests for education, industry, and the military to order and differentiate between individuals[52]. Accordingly, psychometric tools for evaluating response scales have strongly emphasized calculating reliability, while validity has received much less attention, and generally, the tools for assessing validity have also been quantitative[55] (see, e.g., ref. 56). Quantification was further reinforced by the historical intertwining of psychology and statistics; for example, statistics also developed through psychological research topics such as general intelligence, in which correlational and factor analysis played a key role[57,58].

As is clear from this history, quantification and statistical analysis are often seen as essential for scientific rigor. In psychological science in particular, the idea of objectivity is seen as essential for research quality and is closely linked to the use of statistical methods[54]. A common belief is that quantification and corresponding statistical techniques inherently reduce personal biases and prevent subjective influences from affecting results[59,60]. This perspective thus assumes that collecting quantitative data, together with the application of statistical procedures, ensures neutrality by minimizing individual intuition or subjective judgment in research. Furthermore, quantification is often thought to increase precision and exactness[60,61], allowing comparisons across individuals and within individuals across time. It is not only possible to conclude that an individual is happy; rather, we can also compare their happiness to that of others, showing who is the least or most happy, or what the average level of happiness is[62]. Practically speaking, quantification makes it easier to summarize data, identify patterns, and generalize findings, especially with datasets containing a large number of data points (per person), as is typical for ESM studies; this is something far more time-consuming to achieve with qualitative data[54,55,60]. Finally, as discussed above, psychometric models used to evaluate construct validity and reliability are generally designed for quantitative data (see also ref. 60). Taken together, these factors incentivize the collection, analysis, and validation of (closed-ended) quantitative data instead of (open-ended) qualitative data[55].

The essential role of qualitative research

The status of qualitative methods in psychology

Despite the dominance and recognized value of quantitative research in psychology, qualitative methods have never been entirely absent[63]. Although pioneers of psychology such as Wilhelm Wundt, Sigmund Freud, and William James affirmed the scientific value of qualitative approaches, qualitative methods remained largely absent from psychology's educational curricula throughout much of the 20th century[63]. As psychology sought to establish itself as an objective science, it was natural to adopt closed-ended scales, such as Likert scales, for self-report data collection (see p. 17 in ref. 36). At the same time, prominent individual researchers such as Gordon Allport[64], Daniel Brower[65], Sigmund Koch[66], and Solomon Asch[67] continued to explicitly advocate for qualitative methods, even as such methods were sometimes labeled as pseudoscientific (see p. 19 in ref. 36).

A more general push for qualitative methods occurred in the 1980s, as interest in narrative inquiry, in which stories or narratives are used to describe experiences, increased[68]. For instance, Jerome Bruner differentiated two modes of thought: the logical-scientific or paradigmatic and the narrative[69]. He argued that these two ways of knowing are distinct and essential in how humans make sense of the world, yet the narrative has been largely neglected in psychological science. In contrast to abstract logical-scientific knowledge (paradigmatic), narrative knowledge captures the story-based logic humans use to understand their experiences. Scholars like Polkinghorne[70] continued this line of work, with lasting influences on narrative traditions in psychology and qualitative research[71–73]. This narrative turn also led to increased interest in daily diary methods, in which participants record and describe their experiences in logs or diaries[36]. In line with these developments, Hurlburt[74,75] introduced the "descriptive experience sampling" method, which was an explicitly qualitative form of ESM, aiming to describe experiences instead of quantifying them.

In recent decades, the status of qualitative methods in psychology has gradually moved from being merely recommended to becoming institutionalized and formally recognized (see also Stecyk et al.[76]). For example, the American Psychological Association has a section (APA Division 5) and a journal (*Qualitative Psychology*) for qualitative research methods[63]. Notably, current standards for evaluating the validity of psychological questionnaires acknowledge that gathering validity evidence often requires both quantitative and qualitative approaches: "Both qualitative and quantitative sources of evidence are important in evaluating whether items are psychometrically sound…" (American Educational Research Association [AERA][77], p. 64; see also Wolf[55]). Despite this, validity evidence remains predominantly quantitative, and forms of validity that require qualitative methods, such as face validity and response process validation (e.g., with cognitive interviewing[78], or event-centered interviewing[79]), are largely absent in meta-analyses summarizing validity evidence in self-report questionnaires. Even in ESM studies that collect and analyze open-ended items, very few use these to check or improve validity, for example, by asking about the response processes (for exceptions, see, e.g., Schorrlepp et al.[26]; see also the Supplementary Note). Additionally, there is a lack of clear standards for how qualitative validity evidence should be collected and reported[55].

Limits of quantitative data in psychology

Close-ended (Likert) scales and quantitative methods are dominating the field of psychology, and are widely believed to produce more rigorous and objective results than qualitative methods (see "The enduring influence of quantification" section). However, this belief may not hold up to scrutiny. Consider, for example, Kahneman's definition of 'objective happiness' as 'the average of utility over a period of time', which reflects a strong belief in objective self-report measurement[59,80]. Yet whether self-report data are collected in qualitative or quantitative form, they originate from the same source: the participant[80]. In both cases, the participant expresses an internal state, either by choosing a number or by using words and sentences. The primary distinction, therefore, is not one of objectivity versus subjectivity[59], but one of abstraction. Saying 'sunlight was streaming through the café window, and I couldn't stop smiling while talking' is less abstract than selecting a 5 on a scale measuring how happy you feel right now. Numbers thus compress and reduce experience, making it more straightforward to compare across individuals, but leaving out much of the complexity and nuance that can be expressed in language[81].

Moreover, quantitative data can give the misleading impression that there are no unanswered questions about the psychological realities they represent[54]. Expressing something in numerical form often implies completeness, that the numbers are sufficient and no further exploration is necessary[54]. Yet, psychological phenomena are rarely so neatly bound. Researchers cannot anticipate every relevant factor in advance, and unexpected experiences or motives remain invisible to predefined numbers or categories. For example, Truijens et al.[82] discuss a case in which a participant started entering random data out of frustration over the fact that the data was not shared with their therapist. Such unexpected occurrences and the reasons behind them cannot be seen in quantitative data alone. Quantification, therefore, restricts researchers to anticipated forms of information and is inherently limited in its ability to detect and explain surprising or novel insights[83].

Furthermore, numbers can give a sense of false precision or accuracy, a phenomenon sometimes referred to as *spurious precision*[61]. A participant might select a score of 6 for sad mood, but not because their experience maps

Box 2 | Reasons for including open-ended items in ESM

1. Understanding response processes
   a. Item and response scale interpretation
   b. Detecting cultural and contextual differences
   c. Response shifts
2. Detecting and understanding anomalies in data, such as missing values or unexpected responses
3. Capturing the complexity and uniqueness of lived experience
4. Understanding complex and context-dependent causal processes (the *why* behind outcomes or patterns)
5. Identifying processes at different time scales and their temporal ordering

exactly onto that number. Rather, it may be because the participant feels that their mood is somewhere between 5 and 8, but not a 4 or a 9, and this nuance is masked because the participant has to select a single number. All of these issues and examples show how numbers may hide underlying problems, reasons, and assumptions, which cannot be detected with quantitative data alone[26,83].

Reasons for including open-ended items in ESM

Due to the limitations of quantitative data, qualitative data are essential to determine the meaning of quantitative responses to self-reports[55,65]. That is, we must understand how participants interpret the items and how they arrive at a specific number or category on a response scale[78]. Currently in ESM studies, researchers often simply assume that participants interpret and respond to a questionnaire as intended, without actually verifying it. In practice, participants can interpret items and response scales quite differently from each other, especially when they come from different cultural or social backgrounds (see, e.g., Scollon et al.[84]). For example, in one study, the item "physical activity" meant different things to different participants. Those who had just arrived in the Netherlands described biking as a physical activity, whereas participants who had lived there longer considered it merely a form of transportation[26]. Similarly, how does a researcher know whether a participant selecting the 'doing nothing' option means lying on the couch while listening to music, or literally doing nothing at all[26,31]? In the same study[26], one participant used 4 as a neutral score, while another used 8 on a scale of 1–10 to express neutrality. Interpretation may also shift over time, a phenomenon known as *response shift*[26,85]. As participants repeatedly complete the same ESM questions, their understanding of emotions or scale endpoints may change. For instance, a day spent alone, not leaving the house, and eating nothing may have originally felt like a 1 (on a happiness scale from 1 to 10), but after the participant experiences a personal loss, this endpoint may seem completely inadequate, as the participant would like to express an even lower score[26,81,85].

Other distortions may result from disengagement, fatigue, or sadness, leading to missing responses or shallow engagement with the items (see, e.g., the example of a participant entering random data mentioned in the previous section). These issues can often be detected only by asking participants, for example, by including open-text boxes in ESM protocols to let participants explain their answers and their context, which can shed light on their interpretation, reasons for skipping entries, or disengagement[16,26,83,86].

A further advantage of using open-ended items is that they provide access to the complexity of lived experience by allowing participants to express themselves in their own words[83,86,87]. For example, Hurlburt, Happé, and Frith used Descriptive Experience Sampling with autistic individuals, and the results indicated that their reported inner experience often consisted only of visual imagery, with little or no inner speech or emotional awareness[88]. This finding had potentially significant implications for autism research, suggesting that the difficulty in perspective-taking experienced by some autistic individuals might stem from the absence of internal verbal or emotional representations[88]. Such findings illustrate what is lost when we rely solely on quantitative data: not just nuance, but often also the core phenomena we aim to understand[86]. In the case of autism, researchers might not have studied internal cognitive styles at all, despite their potential to offer essential insights into autistic experience and development.

As the cases discussed above show, open-ended responses are also helpful when it comes to understanding causality, in particular, the *why* behind certain outcomes, patterns, or anomalies in data, such as missing values or unexpected responses[83]. Consider, for example, predefined categories for social context and activities. Participants may report, on average, being happier when they are with someone (e.g., "with a friend") and engaged in "self-care" (another predefined category). However, you may find that a participant who selected these categories still reports being unhappy. An open-text box can help clarify the reasons for this apparent discrepancy: *"Friend was doing my hair but burned my skin with the hair curler and I was in tremendous pain. Furthermore, I just do not have a buffer anymore since the financial problems started last month and they are still stressing me"* (adjusted example from Stadel et al.[89]). Understanding the unique momentary context and how the person came to such a low score on happiness would have been impossible to derive from predefined categories and the resulting quantitative data alone. Quantitative data are suited for studying increases or decreases in, for example, emotions or behaviors, whereas qualitative data can complement these analyses by offering insights into complex and context-dependent causal processes[83,90].

Furthermore, processes that unfold across different time scales, such as financial problems that occurred last month but are still having an impact on a day-to-day basis, are difficult to capture through quantitative measures alone. Due to their less constrained format, open-ended items in ESM can allow researchers to identify start- and endpoints of activities or processes operating on longer time scales (e.g., a participant describing traveling to another country by plane and then a few days later mentioning the return flight[89]). Additionally, open-ended descriptions can contain information on the temporal order of events (e.g., "*after volunteering at [a café] i came back home and took a nap because i was too tired. then after waking up i worked on my thesis, called my dad and now i'm about to make dinner*"[89]) or refer to earlier events to provide more context (see the example about financial problems in the previous paragraph).

To conclude, quantitative or categorical ESM items alone cannot fully capture the details and context of daily life. Quantitative data only skim the surface of what people actually experience, and it would not be feasible to list every possible category in advance, as the range of experiences people have is vast and varied[91]. As a result, it is unlikely that quantitative ESM data alone can explain why people felt or acted in a particular way[83]. Instead of trying to capture the vast complexity of different time scales or attempting to account for every possible scenario with numbers and predefined categories, it is more effective to collect qualitative data and build understanding from there[74,90] (see Box 2 for a summary of the reasons for including open-ended items).

The future of open-ended items in ESM

ESM research initially included both closed- and open-ended items, but the field has increasingly focused on the former. This dominance of quantitative data is understandable: numbers are easily comparable, patterns can be quickly summarized, and there are well-established methods to analyze the data and evaluate psychometric quality. In line with this, open-ended responses are rarely analyzed or reported, even though many studies still collect them. In general, there is little explicit discussion of the rationale for including or excluding open-ended items, or, when included, why they are not further analyzed or used. One possible contributing factor is the widespread tendency to follow established ESM designs, often rooted in early work such as Delespaul[92], which may lead researchers to include open-ended items more out of convention than active intention to analyze them. However, there has been a notable shift in recent years in the literature, as the proportion of ESM papers that actually analyze open-ended items has been increasing. Although this is still a very small proportion of all ESM studies, it nevertheless suggests that interest in open-ended items is growing. This brings us back to a central question: should open-ended items be set aside due to the effort they require, or is it time to, at a large scale, collect, code, and analyze them as an integral part of ESM research?

We think the answer to the latter question is a clear "yes": It is exactly because of the intrinsic limitations of quantitative data that qualitative or open-ended items are essential in ESM. The abstraction inherent in quantitative data should not be mistaken for scientific rigor or objectivity. In the end, both quantitative and qualitative data come from the same source, humans reporting on their experiences. Quantitative data do not provide a more scientific representation of reality; they are simply representations where the richness of experience has been filtered into numbers and predefined categories. In this way, the upside of quantitative data becomes its biggest downside: once data are abstracted, we no longer know how they came to be. Moving too quickly to quantification, or failing to allow space for revision or new information, risks producing "false abstractions"[36,64] and leads away from life as it is lived, from human experiences in all their complexity[36,93]. Therefore, careful qualitative description should serve as the foundation on which quantification is built[36,74,93], as already suggested by early proponents of qualitative methods such as Allport[64], Brower[65], Koch[66], and Asch[67].

When researchers also have access to the qualitative data in open-ended responses, they can furthermore perform the abstraction themselves, making the process more transparent and standardized. Researchers can then document and share how they arrived at particular categories or interpretations, enabling others to assess, question, or reproduce their reasoning (for a discussion, see Braun & Clarke[94]). This approach opens up the possibility of evaluating both the participant's abstraction (e.g., the number they gave or the category they chose) and the researcher's, making the interpretive process more transparent and subject to scrutiny, something that is rarely possible with purely quantitative data. In this way, incorporating open-ended qualitative data actually enhances the scientific rigor of ESM studies (see also Douglas[95] and Kidder & Fine[96]). Moreover, retaining access to different levels of abstraction allows researchers to revisit the data with new research questions in mind.

Qualitative data are also indispensable for addressing aspects of experiences that are hard to quantify: the "why" behind responses, unexpected patterns, and the multiple time scales at which experiences and events unfold[83,90,97]. Numerical abstractions are not sufficient for capturing these dimensions. For example, what looks like a stable pattern in the numbers may turn out to mask essential differences. Two participants might both select "3" on a happiness scale from 1 to 5, but in the open-text items, one might describe their mood as "peaceful and content after a walk in nature", while the other felt "emotionally numb but not actively distressed". These are very different experiences, which would not be visible from the quantitative data alone. Some form of open-ended items might therefore always be necessary to improve the validity of the quantitative data, to capture the unexpected, and to study patterns, reasons, and phenomena that resist quantification. Thus, we strongly advocate for mixed methods in ESM, collecting qualitative data alongside quantitative data and using them together to enrich the insights drawn from ESM studies[91,96–100].

Suggestions for future research

While the added value of open-ended items is clear, it is less clear how to get the most out of such data, both in terms of collection and analysis. One important opportunity that remains underexplored in the ESM field is how to formulate questions and guide participants toward providing the right level of detail in their responses. As noted by Hektner et al.[11], individuals vary in how they describe their daily activities or experiences: some offer detailed accounts, while others remain abstract or vague. Importantly, item wordings should encourage descriptions that are of appropriate length for further qualitative or quantitative analyses, such as natural language processing techniques (e.g., Boyd[101]). Hektner et al.[11] recommend that researchers define their research questions and coding categories in advance whenever possible, so that participants can be appropriately instructed from the start. For example, participants could be encouraged, depending on the study aims, to avoid abstract statements (e.g., "I feel stressed"), to explain the reasons for their feelings, to provide contextual information, or simply to give more details in their answers (see also Fritz et al.[102]).

In addition, including open-ended items alongside predefined activity categories can also be beneficial. This not only allows researchers to see how participants would categorize their own activities, but also helps capture more routine behaviors, such as showering, using the bathroom, or brushing teeth, which participants may overlook or deem too mundane to mention in open-text responses[89]. A combination of predefined categories and open-ended items, therefore, ensures a richer contextual picture without forcing researchers to rely solely on participants' self-categorization. One potential challenge with this is that participants might also adopt the abstract, predefined categories to describe their experiences in the open-ended items, thereby limiting the richness and spontaneity of their responses. However, preliminary results[89] suggest that participants do not tend to systematically reduce their responses to overly abstract categories, indicating that this challenge is likely to be manageable in practice.

It is also important to keep in mind that completing open-ended items to obtain qualitative data takes more time than simply responding to Likert-type or visual analog scale items. Fortunately, this burden is partly offset by the fact that many participants value the opportunity to express themselves more freely. Although more systematic research is needed, in many studies where participants received feedback on their qualitative data (e.g., word clouds), compliance was high, and participants continued to provide open-ended responses over an extended time period[89,97,103–105]. Also, new data collection techniques offer further possibilities for alleviating participant burden, although it should be taken into account that they will result in different kinds of information. For example, voice recordings that are subsequently transcribed into text are perceived by many participants as more convenient than open-ended text items, and tend to result in longer responses[97]. Similarly, participants can create short videos or other multimedia content, which can then be transcribed and coded[86]. ESM measures can also be combined with more passive sensing approaches, such as the Electronically Activated Recorder (EAR), which samples brief snippets of daily life[106], or by tracking what participants are typing on their smartphones in different apps throughout the day[107]. The latter especially will result in qualitative data that can complement quantitative ESM data, but will contain different information compared to prompted open-ended items.

Regarding data analysis, given that ESM datasets can be very large, including many participants and up to hundreds of time points, it is also important to look into the new possibilities offered by automated text-analytic tools, such as large language models (LLMs) and other artificial intelligence tools (AI). Such tools can help to lift the burden in analyzing large datasets, for example, by anonymizing[108] and coding[22,109] qualitative responses into predefined categories[110]. However, it is crucial to approach the use of AI with care (e.g., not simply sharing data with online LLMs owned by private corporations), as well as to keep in mind that the choice of analytic tools must be driven by the research goals and study-specific considerations.

Another important challenge is that ESM researchers often lack the tools and knowledge to effectively analyze open-ended data. Currently, the analysis of open-ended ESM data typically involves coding and counting participants' categorizations (i.e., applying content analysis, see Hsieh & Shannon[111]), and then applying inferential statistics on them (see Table 1 and Supplementary Note). For example, Delespaul et al.[112] collected open-ended responses on the context and activities of schizophrenic vs. non-schizophrenic individuals, and then coded this data into categories to use it

Box 3 | Topics for future research on open-ended items in ESM

1. Studying what constitutes a high-quality open-ended item and how to formulate open-ended items in ways that elicit the appropriate level of detail and context from participants
2. Determining when and how to combine predefined categories with open-text responses to capture both routine patterns and richer contextual information
3. Studying how participant burden can be reduced while still allowing for rich qualitative data collection, for example, using voice recordings
4. Increasing the efficiency and feasibility of transcribing and (quantitatively) analyzing open-ended responses, for example, using automated text-analytic tools or large language models (LLMs)
5. Implementing qualitative approaches within ESM, such as reflexive thematic analysis or interpretative phenomenological analysis
6. Integrating qualitative data as a standard component of the iterative ESM research cycle, from conceptualizing constructs to evaluating numerical measures
7. Developing concrete tutorials, workshops, and including dedicated sections in ESM handbooks on designing and (qualitatively) analyzing open-ended ESM items

in a (multilevel) regression model. While this practice aligns with the overall quantitative orientation of ESM, more advanced qualitative analysis methods for open-ended responses should be considered in future ESM studies, including researchers with experience in these methods as collaborators. This includes qualitative methods such as reflexive thematic analysis[113] or interpretative phenomenological analysis[114]. For example, the work of Sullivan et al.[115] demonstrates how rich insights can be gained from qualitatively analyzing open-ended ESM responses: They collected open-ended descriptions of oppression-based stress experiences in daily life among sexual and gender minority adolescents (SGMA). An inductive content analysis (i.e., not starting from predefined categories) yielded a map of stressors clarifying how both threats and absent safety signals shape SGMA adolescents' daily lives. As another example, De Smet et al.[99] used thematic analyses to analyze open-ended momentary and end-of-day responses together with semi-structured interviews to contextualize quantitative ESM data in depressed youth.

A further recommendation and aim for future research is to make open-ended and qualitative data more front and center in endeavors to improve and check quantitative ESM data. After all, how can we refine our measurement if we do not engage with the entire iterative research cycle: conceptualizing the phenomenon, collecting and analyzing data, and evaluating whether our quantitative measures capture the phenomenon adequately[99,116,117]? Moreover, qualitative data are not only crucial for validity, but also for improving reliability, especially in ESM, where we must consider whether participants interpret and respond to questionnaires consistently over many time points[16]. Currently, open-ended items are only rarely used to check or improve the validity or reliability of the data. Thus, to enhance the quality of ESM data, these kinds of items and other qualitative input should be integrated into ESM studies, so that we can go back and forth between the qualitative data and the numbers, enabling iterative improvement of our measurement[26,99,116] (see also Eisele et al.[118], Eronen[117], Grunfeld et al.[86], and Truijens et al.[119]). For example, open-ended follow-up questions can help clarify which time frame respondents actually use when rating "momentary" emotions or thoughts (e.g., whether they base this on the immediate present, events from the recent past, anticipated future states, or a combination). This kind of information can then be used to adjust item wordings to better capture the intended construct.

Finally, to turn these ideas (see Box 3 for a summary) into reality rather than leaving them as an aspiration, papers like this are not enough. They need to be complemented with overviews and concrete examples of how qualitative data can be used in ESM, such as in Schorrlepp et al.[26] and De Smet et al.[99]. There also remains a lack of resources on more advanced topics such as developing and analyzing open-ended ESM items for qualitative research purposes, for instance phenomenological approaches or applying AI tools (e.g., LLMs) to open-ended ESM data. What is further needed are workshops and tutorials, and ideally, formal education on these topics within psychology curricula, starting already at the bachelor's level, where qualitative methods are underrepresented[116]. In addition, ESM handbooks and methodological guidelines should include dedicated sections on collecting and analyzing open-ended items in ESM research. Taken together, these steps can provide the foundation to systematically integrate open-ended items and qualitative methods into ESM research.

Conclusion

We strongly advocate for the collection and analysis of open-ended items in ESM research, and we hope that handbooks and guidelines on ESM will, in the future, no longer give insufficient attention to them but have dedicated sections on collecting and analyzing open-ended data. Thus, open-ended items should not be included just out of tradition or habit, but should actually be used to improve ESM research and ground our research in the richness of real-world experience, bringing us closer to the phenomena as they are experienced in everyday life.

Received: 16 July 2025; Accepted: 29 January 2026;
Published online: 23 February 2026

References

1. Czsikszentmihalyi, M. & Larson, R. Validity and reliability of the experience sample method. *J. Nerv. Ment. Dis.* **175**, 526–536 (1987).
2. Shiffman, S., Stone, A. A. & Hufford, M. R. Ecological momentary assessment. *Annu. Rev. Clin. Psychol.* **4**, 1–32 (2008).
3. Stone, A. A. & Shiffman, S. Ecological momentary assessment (EMA) in behavioral medicine. *Ann. Behav. Med.* **16**, 199–202 (1994).
4. Trull, T. J. & Ebner-Priemer, U. Ambulatory assessment. *Annu. Rev. Clin. Psychol.* **9**, 151–176 (2013).
5. Fritz, J. et al. So you want to do ESM? 10 Essential topics for implementing the experience-sampling method. *Adv. Methods Pract. Psychol. Sci.* **7**, 25152459241267912 (2024).
6. Kashdan, T. B., Goodman, F. R., Mallard, T. T. & DeWall, C. N. What triggers anger in everyday life? Links to the intensity, control, and regulation of these emotions, and personality traits. *J. Pers.* **84**, 737–749 (2016).
7. Skimina, E. et al. A categorization of behaviors reported in experience sampling studies. *Soc. Psychol. Bull.* **15**, e3029 (2020).
8. Mehl, M. R., & Conner, T. S. (eds.) *Handbook of Research Methods for Studying Daily Life* (Guilford Press, New York, NY, 2012).
9. Myin-Germeys, I. & Kuppens, P. (eds.) *The Open Handbook of Experience Sampling Methodology: A Step-by-Step Guide to Designing, Conducting, and Analyzing ESM Studies* (Center for Research on Experience Sampling and Ambulatory Methods Leuven, Leuven, 2022).
10. Brough, P. *Advanced Research Methods for Applied Psychology: Design, Analysis and Reporting* (Routledge, Abingdon, 2019).

11. Hektner, J. M., Schmidt, J. A. & Csikszentmihalyi, M. *Experience Sampling Method: Measuring the Quality of Everyday Life* (Sage Publications, Thousand Oaks, CA, USA, 2007).
12. Bolger, N., Davis, A. & Rafaeli, E. Diary methods: capturing life as it is lived. *Annu. Rev. Psychol.* **54**, 579–616 (2003).
13. Bolger, N. & Laurenceau, J.-P. *Intensive Longitudinal Methods: An Introduction to Diary and Experience Sampling Research* (Guilford Press, New York, NY, 2013).
14. Trull, T. J. & Ebner-Priemer, U. W. Introduction to the Special Section. *J. Educ. Psychol.* **82**, 183–188 (2009).
15. Trull, T. J. & Ebner-Priemer, U. W. Ambulatory assessment in psychopathology research: a review of recommended reporting guidelines and current practices. *J. Abnorm. Psychol.* **129**, 56–63 (2020).
16. Stone, A. A., Schneider, S. & Smyth, J. M. Evaluation of pressing issues in ecological momentary assessment. *Annu. Rev. Clin. Psychol.* **19**, 107–131 (2023).
17. Pow, J., Lee-Baggley, D. & DeLongis, A. Threats to communion and agency mediate associations between stressor type and daily coping. *Anxiety Stress Coping* **29**, 660–672 (2016).
18. Kirtley, O. J. et al. Initial cohort characteristics and protocol for SIGMA: an accelerated longitudinal study of environmental factors, inter- and intrapersonal processes, and mental health in adolescence. Preprint at https://doi.org/10.31234/osf.io/jp2fk (2021).
19. Weintraub, J., Cassell, D. & DePatie, T. P. Nudging flow through 'SMART' goal setting to decrease stress, increase engagement, and increase performance at work. *J. Occup. Organ. Psychol.* **94**, 230–258 (2021).
20. Jackson, S. D., Mohr, J. J. & Kindahl, A. M. Intersectional experiences: a mixed methods experience sampling approach to studying an elusive phenomenon. *J. Couns. Psychol.* **68**, 299–315 (2021).
21. deVries, M. W. & Delespaul, P. A. E. G. Time, context, and subjective experiences in schizophrenia. *Schizophr. Bull.* **15**, 233–244 (1989).
22. Hoemann, K. et al. Using freely generated labels instead of rating scales to assess emotion in everyday life. *Assessment* **32**, 859–877 (2025).
23. Andrewes, H. E., Hulbert, C., Cotton, S. M., Betts, J. & Chanen, A. M. Ecological momentary assessment of nonsuicidal self-injury in youth with borderline personality disorder. *Personal. Disord.* **8**, 357–365 (2017).
24. Cutler, S., Doherty, S. & Carmichael, B. The experience sampling method: examining its use and potential in tourist experience research. *Curr. Issues Tour.* **21**, 1052–1074 (2018).
25. McCombie, C., Ouzzane, H., Schmidt, U. & Lawrence, V. 'Physically it was fine, I'd eat what normal people do. But it's never like this in my head': a qualitative diary study of daily experiences of life in recovery from an eating disorder. *Eur. Eat. Disord. Rev.* **32**, 46–55 (2024).
26. Schorrlepp, L., Stadel, M., Bringmann, L. F., Hesselink, M. & Maciejewski, D. Utilizing qualitative methods to detect validity issues in clinical experience sampling methodology (ESM). *Psychol. Assess.* **37**, 599–613 (2025).
27. Doherty, S. T., Lemieux, C. J. & Canally, C. Tracking human activity and well-being in natural environments using wearable sensors and experience sampling. *Soc. Sci. Med.* **106**, 83–92 (2014).
28. König, L. M. & Renner, B. Colourful = healthy? Exploring meal colour variety and its relation to food consumption. *Food Qual. Prefer.* **64**, 66–71 (2018).
29. König, L. M. et al. Investigating the relationship between perceived meal colour variety and food intake across meal types in a smartphone-based ecological momentary assessment. *Nutrients* **13**, 755 (2021).
30. Mestdagh, M. et al. m-Path: an easy-to-use and highly tailorable platform for ecological momentary assessment and intervention in behavioral research and clinical practice. *Front. Digit. Health* **5**, https://doi.org/10.3389/fdgth.2023.1182175 (2023).
31. Stadel, M. et al. Capturing the social life of a person by integrating experience-sampling methodology and personal-social-network assessments. *Adv. Methods Pract. Psychol. Sci.* **7**, 25152459241244745 (2024).
32. Rosengaard, L. O., Andersen, M. Z., Rosenberg, J. & Fonnes, S. Several methods for assessing research waste in reviews with a systematic search: a scoping review. *PeerJ* **12**, e18466 (2024).
33. Eisele, G. et al. The effects of sampling frequency and questionnaire length on perceived burden, compliance, and careless responding in experience sampling data in a student population. *Assessment* **29**, 136–151 (2022).
34. Hasselhorn, K., Ottenstein, C. & Lischetzke, T. The effects of assessment intensity on participant burden, compliance, within-person variance, and within-person relationships in ambulatory assessment. *Behav. Res. Methods* **54**, 1541–1558 (2022).
35. Williams, M. T. et al. Compliance with mobile ecological momentary assessment of self-reported health-related behaviors and psychological constructs in adults: systematic review and meta-analysis. *J. Med. Internet Res.* **23**, e17023 (2021).
36. Hyers, L. L. *Diary Methods: Understanding Qualitative Research* (Oxford University Press, Oxford, 2018).
37. Ariens, S., Ceulemans, E. & Adolf, J. K. Time series analysis of intensive longitudinal data in psychosomatic research: a methodological overview. *J. Psychosom. Res.* **137**, 110191 (2020).
38. Bringmann, L. F. et al. A network approach to psychopathology: new insights into clinical longitudinal data. *PLoS ONE* **8**, e60188 (2013).
39. Hamaker, E. L. & Dolan, C. V. Idiographic data analysis: quantitative methods from simple to advanced. In *Dynamic Process Methodology in the Social and Developmental Sciences* (eds Valsiner, J. et al.) 191–216 (Springer-Verlag, New York, NY, 2009).
40. Schuurman, N. K., Ferrer, E., de Boer-Sonnenschein, M. & Hamaker, E. L. How to compare cross-lagged associations in a multilevel autoregressive model. *Psychol. Methods* **21**, 206–261 (2016).
41. Kuppens, P., Oravecz, Z. & Tuerlinckx, F. Feelings change: accounting for individual differences in the temporal dynamics of affect. *J. Pers. Soc. Psychol.* **99**, 1042–1060 (2010).
42. Suls, J., Green, P. & Hillis, S. Emotional reactivity to everyday problems, affective inertia, and neuroticism. *Pers. Soc. Psychol. Bull.* **24**, 127–136 (1998).
43. Chung, J. M., Harari, G. M. & Denissen, J. J. A. Investigating the within-person structure and correlates of emotional experiences in everyday life using an emotion family approach. *J. Pers. Soc. Psychol.* **122**, 1146–1189 (2022).
44. Kuppens, P., Dejonckheere, E., Kalokerinos, E. K. & Koval, P. Some recommendations on the use of daily life methods in affective science. *Affect. Sci.* **3**, 505–515 (2022).
45. Horstmann, K. T., Rauthmann, J. F., Sherman, R. A. & Ziegler, M. Distinguishing simple and residual consistencies in functionally equivalent and non-equivalent situations: evidence from experimental and observational longitudinal data. *Eur. J. Personal.* **35**, 833–860 (2021).
46. Schuurman, N. K. & Hamaker, E. L. Measurement error and person-specific reliability in multilevel autoregressive modeling. *Psychol. Methods* **24**, 70–91 (2019).
47. Castro-Alvarez, S., Tendeiro, J. N., de Jonge, P., Meijer, R. R. & Bringmann, L. F. Mixed-effects trait-state-occasion model: studying the psychometric properties and the person–situation interactions of psychological dynamics. *Struct. Equ. Model. Multidiscip. J.* **29**, 438–451 (2022).
48. Castro-Alvarez, S., Bringmann, L. F., Back, J. & Liu, S. The many reliabilities of psychological dynamics: an overview of statistical approaches to estimate the internal consistency reliability of

intensive longitudinal data. *Psychol. Methods* https://doi.org/10.1037/met0000778 (2025).
49. Eisele, G., Kasanova, Z. & Houben, M. Questionnaire design and evaluation. In *Experience sampling methodology: The open handbook of a step-by-step guide to designing, conducting, and analyzing ESM studies* 71–90 (The Center for Research on Experience Sampling and Ambulatory Methods, Leuven, 2022).
50. Haslbeck, J. M. B. et al. Comparing Likert and visual analogue scales in ecological momentary assessment. *Behav. Res. Methods* **57**, 217 (2025).
51. Wright, A. G. C., Ringwald, W. R. & Zimmermann, J. Measuring psychopathology in daily life. *Clin. Psychol. Sci.* **13**, 649–663 (2025).
52. Michell, J. Normal science, pathological science and psychometrics. *Theory Psychol* **10**, 639–667 (2000).
53. Porter, T. *Trust in Numbers* (Princeton University Press, 1995).
54. Danziger, K. *Constructing the Subject: Historical Origins of Psychological Research* (Cambridge Univ. Press, Cambridge, 1990).
55. Wolf, M. G. The problem with over-relying on quantitative evidence of validity. Preprint at https://doi.org/10.31234/osf.io/v4nb2 (2023).
56. Forkmann, T. et al. Assessing suicidality in real time: a psychometric evaluation of self-report items for the assessment of suicidal ideation and its proximal risk factors using ecological momentary assessments. *J. Abnorm. Psychol.* **127**, 758–769 (2018).
57. Eronen, M. I. & Romeijn, J.-W. Philosophy of science and the formalization of psychological theory. *Theory Psychol* **30**, 786–799 (2020).
58. Revelle, W., Wilt, J. & Condon, D. M. Individual differences and differential psychology: A brief history and prospect. In *The Wiley-Blackwell handbook of individual differences* 3–38 (Wiley Blackwell, 2011).
59. Gelman, A. & Hennig, C. Beyond subjective and objective in statistics. *J. R. Stat. Soc. Ser. A Stat. Soc.* **180**, 967–1033 (2017).
60. Wijsen, L. D., Borsboom, D. & Alexandrova, A. Values in psychometrics. *Perspect. Psychol. Sci.* **17**, 788–804 (2022).
61. Newfield, C., Alexandrova, A. & John, S. *Limits of the Numerical: The Abuses and Uses of Quantification* (The University of Chicago Press, 2022).
62. Alexandrova, A. *A Philosophy for the Science of Well-Being* (Oxford University Press, 2017).
63. Wertz, F. J. Qualitative inquiry in the history of psychology. *Qual. Psychol.* **1**, 4–16 (2014).
64. Allport, G. W. *The Use of Personal Documents in Psychological Science: Prepared for the Committee on Appraisal of Research*. xix, 210 (Social Science Research Council, New York, NY, US, 1942)
65. Brower, D. The problem of quantification in psychological science. *Psychol. Rev.* **56**, 325–333 (1949).
66. Koch, S. Epilogue. in *Psychology: A Study of a Science. Volume 3. Formulations of the Person and the Social Context* (ed. Koch, S.) 729–788 (McGraw-Hill, 1959).
67. Asch, S. A perspective on social psychology. In *Psychology: A Study of a Science. Volume 3. Formulations of the Person and the Social Context* (ed. Koch, S.) 363–383 (McGraw-Hill, 1959).
68. Polkinghorne, D. E. Narrative configuration in qualitative analysis. *Int. J. Qual. Stud. Educ.* **8**, 5–23 (1995).
69. Bruner, J. Chapter VI: narrative and paradigmatic modes of thought. *Teach. Coll. Rec.* **86**, 97–115 (1985).
70. Polkinghorne, D. E. *Narrative Knowing and the Human Sciences*. xi, 232 (State University of New York Press, Albany, NY, USA, 1988).
71. Monteagudo, J. G. Jerome Bruner and the challenges of the narrative turn: then and now. *Narrat. Inq.* **21**, 295–302 (2011).
72. Josselson, R. & Hammack, P. L. *Essentials of Narrative Analysis*. viii, 102 (American Psychological Association, Washington, DC, USA, 2021).
73. McLeod, J. *Narrative and Psychotherapy*. xi, 180 (Sage Publications, Inc, Thousand Oaks, CA, USA, 1997).
74. Hurlburt, R. T. Randomly sampling thinking in the natural environment. *J. Consult. Clin. Psychol.* **65**, 941–949 (1997).
75. Hurlburt, R. T. *Sampling Normal and Schizophrenic Inner Experience*. xvii, 287 (Plenum Press, New York, NY, USA, 1990).
76. Stecyk, T., Wendt, D. C. & Blackmore, S. Publication trends for qualitative inquiry in American Psychological Association and Association for Psychological Science journals. *Am. Psychol.* https://doi.org/10.1037/amp0001526 (2025).
77. AERA, APA, & NCME. *Standards for Educational and Psychological Testing* (Author, American Educational Research Association, Washington, DC, 2014).
78. Willis, G. *Cognitive Interviewing* (Sage Publications, 2005).
79. Kelly, B. C. & Sennott, C. Event-centered interviewing: integrating qualitative interviews with experience sampling technologies. *Sociol. Methodol.* **55**, 1–24 (2025).
80. Kahneman, D. Objective happiness. In *Well-being: The Foundations of Hedonic Psychology* 3–25 (Russell Sage Foundation, New York, NY, USA, 1999).
81. Truijens, F. L. et al. When quantitative measures become a qualitative storybook: a phenomenological case analysis of validity and performativity of questionnaire administration in psychotherapy research. *Qual. Res. Psychol.* **19**, 244–287 (2022).
82. Truijens, F. L., De Smet, M. M., Vandevoorde, M., Desmet, M. & Meganck, R. What is it like to be the object of research? On meaning making in self-report measurement and validity of data in psychotherapy research. *Methods Psychol* **8**, 100118 (2023).
83. De Smet, M. M., Schoofs, M., Peeters, H., Van Nieuwenhove, K. & Meganck, R. Toward meaningful networks: how qualitative research can inform idiographic assessment and personalized care. *Qual. Psychol*. https://doi.org/10.1037/qup0000314 (2024)
84. Scollon, C. N., Koh, S. & Au, E. W. M. Cultural differences in the subjective experience of emotion: when and why they occur. *Soc. Personal. Psychol. Compass* **5**, 853–864 (2011).
85. Schwartz, C. E. & Rapkin, B. D. Reconsidering the psychometrics of quality of life assessment in light of response shift and appraisal. *Health Qual. Life Outcomes* **2**, 16 (2004).
86. Grunfeld, G., Bringmann, L. F. & Fulford, D. Putting the “experience” back in experience sampling: a phenomenological approach. *J. Psychopathol. Clin. Sci*. https://doi.org/10.1037/abn0000928 (2024).
87. Hesse-Biber, S. Qualitative approaches to mixed methods practice. *Qual. Inq.* **16**, 455–468 (2010).
88. Hurlburt, R. T., Happé, F. & Frith, U. Sampling the form of inner experience in three adults with Asperger syndrome. *Psychol. Med.* **24**, 385–395 (1994).
89. Stadel, M., Langener, A. M., Hoemann, K. & Bringmann, L. F. Assessing daily life activities with experience sampling methodology (ESM): scoring predefined categories or qualitative analysis of open-ended responses? *Methods Psychol* **12**, 100177 (2025).
90. Maxwell, J. A. & Levitt, H. M. How qualitative methods advance the study of causation in psychotherapy research. *Psychother. Res.* **33**, 1019–1030 (2023).
91. von Klipstein, L. et al. Opening the contextual black box: a case for idiographic experience sampling of context for clinical applications. *Qual. Life Res*. https://doi.org/10.1007/s11136-024-03848-0 (2024).
92. Delespaul, P. A. E. G. *Assessing Schizophrenia in Daily Life: The Experience Sampling Method*. Doctoral thesis, Maastricht Univ. https://doi.org/10.26481/dis.19950504pd (1995).
93. Rozin, P. Social psychology and science: some lessons from Solomon Asch. *Personal. Soc. Psychol. Rev.* **5**, 2–14 (2001).
94. Braun, V. & Clarke, V. Conceptual and design thinking for thematic analysis. *Qual. Psychol.* **9**, 3–26 (2022).
95. Douglas, H. *Science, Policy, and the Value-Free Ideal* (Univ. Pittsburgh Press, 2009).

96. Kidder, L. H. & Fine, M. Qualitative and quantitative methods: when stories converge. *New Dir. Program Eval.* **1987**, 57–75 (1987).
97. Kaplan, D. M. et al. Fabla: A voice-based ecological assessment method for securely collecting spoken responses to researcher questions. *Behav. Res. Methods* **57**, 257 (2025).
98. De Smet, M. M. et al. What "good outcome" means to patients: understanding recovery and improvement in psychotherapy for major depression from a mixed-methods perspective. *J. Couns. Psychol.* **67**, 25–39 (2020).
99. De Smet, M. M., Quintens, L. I., Debeer, D. & Jongerling, J. Adding meaning and context to nodes and edges: a mixed methods approach to enhance the validity of idiographic networks. *Cogn. Ther. Res*. https://doi.org/10.1007/s10608-025-10670-6 (2025).
100. Östlund, U., Kidd, L., Wengström, Y. & Rowa-Dewar, N. Combining qualitative and quantitative research within mixed method research designs: a methodological review. *Int. J. Nurs. Stud.* **48**, 369–383 (2011).
101. Boyd, R. L. Psychological text analysis in the digital humanities. In *Data Analytics in Digital Humanities* (ed. Hai-Jew, S.) 161–189 (Springer International Publishing, Cham, 2017).
102. Fritz, J. et al. *Idiographic Patient Reported Outcome Measures (I-PROMs): A Narrative Systematic Review on Patient-Centred Measurement in Psychotherapy* https://doi.org/10.31234/osf.io/7jqkg_v1 (2025).
103. Bringmann, L. F. et al. Developing a qualitative and quantitative ambulatory assessment-based feedback system within cognitive behavioural interventions for people with persecutory beliefs. *Internet Interv.* **40**, 100819 (2025).
104. Bringmann, L. F., van der Veen, D. C., Wichers, M., Riese, H. & Stulp, G. ESMvis: a tool for visualizing individual experience sampling method (ESM) data. *Qual. Life Res.* **30**, 3179–3188 (2021).
105. von Klipstein, L., Servaas, M. N., Schoevers, R. A., van der Veen, D. C. & Riese, H. Integrating personalized experience sampling in psychotherapy: a case illustration of the Therap-i module. *Heliyon* **9**, e14507 (2023).
106. Mehl, M. R. The Electronically Activated Recorder (EAR): a method for the naturalistic observation of daily social behavior. *Curr. Dir. Psychol. Sci.* **26**, 184–190 (2017).
107. Bemmann, F. et al. Putting language into context using smartphone-based keyboard logging. Preprint at https://doi.org/10.48550/arXiv.2403.05180 (2024).
108. Zuo, Z. et al. Data Anonymization for Pervasive Health Care: Systematic Literature Mapping Study. *JMIR Med. Inform.* **9**, e29871 (2021).
109. Tov, W., Ng, K. L., Lin, H. & Qiu, L. Detecting well-being via computerized content analysis of brief diary entries. *Psychol. Assess.* **25**, 1069–1078 (2013).
110. Feuerriegel, S. et al. Using natural language processing to analyse text data in behavioural science. *Nat. Rev. Psychol.* **4**, 96–111 (2025).
111. Hsieh, H. F. & Shannon, S. E. Three approaches to qualitative content analysis. *Qual. Health Res.* **15**, 1277–1288 (2005).
112. Delespaul, P., deVries, M. & Van Os, J. Determinants of occurrence and recovery from hallucinations in daily life. *Soc. Psychiatry Psychiatr. Epidemiol.* **37**, 97–104 (2002).
113. Braun, V. & Clarke, V. Using thematic analysis in psychology. *Qual. Res. Psychol.* **3**, 77–101 (2006).
114. Smith, J. A. Interpretative phenomenological analysis: getting at lived experience. *J. Posit. Psychol.* **12**, 303–304 (2017).
115. Sullivan, T. R., Flynn, S. S., Touhey, S. & Mereish, E. H. Through their eyes: a qualitative, daily diary exploration of oppression-based stress experiences among sexual and gender minority adolescents. *Psychol. Sex. Orientat. Gend. Divers*. https://doi.org/10.1037/sgd0000779 (2024).
116. Bringmann, L. F., Elmer, T. & Eronen, M. I. Back to basics: the importance of conceptual clarification in psychological science. *Curr. Dir. Psychol. Sci.* **31**, 340–346 (2022).
117. Eronen, M. I. Causal complexity and psychological measurement. *Philos. Psychol.* **38**, 2217–2232 (2025).
118. Eisele, G. et al. ESM-Q: A consensus-based quality assessment tool for experience sampling method items. *Behav. Res. Methods* **57**, 124 (2025).
119. Truijens, F. L., Cornelis, S., Desmet, M., De Smet, M. M. & Meganck, R. Validity beyond measurement: why psychometric validity is insufficient for valid psychotherapy research. *Front. Psychol*. **10**, https://doi.org/10.3389/fpsyg.2019.00532 (2019).
120. Shin, D., Kim, H., Lee, S., Cho, Y. & Jung, W. Using large language models to detect depression from user-generated diary text data as a novel approach in digital mental health screening: instrument validation study. J. Med. Internet Res. 26, e54617 (2024).
121. Rauschenberg, C. et al. Stress sensitivity as a putative mechanism linking childhood trauma and psychopathology in youth's daily life. *Acta Psychiatr. Scand.* **136**, 373–388 (2017).
122. Farmer, S., Mindry, D., Comulada, S. W. & Swendeman, D. Mobile phone ecological momentary assessment of daily stressors among people living with HIV: elucidating factors underlying health-related challenges in daily routines. *J. Assoc. Nurses AIDS Care* **28**, 737–751 (2017).
123. Flynn, S.S., Touhey, S., Sullivan, T.R., & Mereish, E.H. Queer and transgender joy: a daily diary qualitative study of positive identity factors among sexual and gender minority adolescents. *Psychol. Sex. Orientat. Gend. Divers.* https://doi.org/10.1037/sgd0000733 (2024).

Acknowledgements

Ideas for this paper originated from discussions at the workshop "It's About Time – Improving Intensive Longitudinal Data" convened by the Measurement Is The New Black (MITNB; mitnb.org) initiative from 4 to 8 November 2024 at the Lorentz Center, Leiden, the Netherlands. Icons in Box 1 were obtained from the Noun Project (thenounproject.com). In addition, we are grateful for discussions with Malin Meyer on time frames, with Annelise Blanchard on the importance of qualitative information, and with Fridtjof Petersen on using intervals instead of single numbers in item responses. These discussions provided an important source of inspiration for this paper. The research of Laura Bringmann was partly funded by an Open Competition M grant of the Dutch Research Council (NWO; grant number 406.22.24GO.006), as well as by the Stress in Action project (https://www.stress-in-action.nl), which is financially supported by the Dutch Research Council and the Dutch Ministry of Education, Culture and Science (NWO gravitation grant number 024.005.010). The research of Melissa De Smet was partly funded by a Veni grant of the Dutch Research Council (NWO; VI.Veni.221 G.103). The research of Charlotte Vrijen was partly funded by a Veni grant of the Dutch Research Council (NWO; grant number VI.Veni.211 G.033). The funders had no role in the preparation of the manuscript or the decision to publish. During the preparation of this work, the authors used ChatGPT for language editing. After using this tool, the authors reviewed and edited the content as needed and take full responsibility for the publication's content.

Author contributions

L.F.B.: Conceptualization, Investigation, Writing – original draft, Writing – review & editing, Project administration. G.R.G.: Conceptualization, Investigation, Writing – original draft, Writing – review & editing. L.S.: Conceptualization, Investigation, Writing – original draft, Writing – review & editing. L.V.D.E.V.: Conceptualization, Investigation, Writing – review & editing. M.M.D.S.: Conceptualization, Investigation, Writing – original draft, Writing – review & editing. C.V.: Conceptualization, Investigation, Writing – original draft, Writing – review & editing. E.R.: Conceptualization, Writing – original draft, Writing – review & editing. J.F.: Conceptualization,

Investigation, Writing – review & editing. J.X.C.: Investigation, Writing – review & editing. O.K.: Conceptualization, Investigation, Writing – review & editing. M.I.E.: Conceptualization, Writing – original draft, Writing – review & editing. J.M.C.: Conceptualization, Investigation, Writing – original draft, Writing – review & editing. M.S.: Conceptualization, Investigation, Visualization, Writing – original draft, Writing – review & editing. K.H.: Conceptualization, Investigation, Writing – original draft, Writing – review & editing.

Competing interests
The authors declare no competing interests.

Additional information
Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s44271-026-00416-9.

Correspondence and requests for materials should be addressed to Laura F. Bringmann.

Peer review information *Communications Psychology* thanks Michael Krämer, Deanna Kaplan, Laura König for their contribution to the peer review of this work. Primary Handling Editor: Jennifer Bellingtier. A peer review file is available.

Reprints and permissions information is available at http://www.nature.com/reprints

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.