



---

# 1

## Introduction

I wrote this volume in the service of two complementary goals. First, I wanted to provide an introduction to the hows and whys of diary-style research for scholars who are interested in conducting research using diary methods. Second, I wanted to provide scholars, who may not be interested in conducting research using diary methods, sufficient background so that they can make sense of articles describing diary research. To accomplish these goals I discuss how to conceptualize and design a diary study, how to conduct a diary study, and how to analyze the data and report the results of these analyses.

Be forewarned, the style of the writing in this volume is not traditionally academic. Although there are a decent number of references, I have written the volume in a style that resembles what I might say to a colleague as we are discussing a study. I have been conducting diary research for close to 40 years, and, during this time, I would like to think that I have accumulated what would be considered a decent amount of experience. One reason I wrote this volume was to share this experience with a wider audience than the colleagues with whom I have worked. At times, I make recommendations based primarily upon my experience or my general impression of the state of affairs. This is because, for many issues, little if any formal research has been done. As a matter of convenience, many of the examples I use or discuss in more detail are from research I have conducted or with which I have been involved in some way. I did this because I am more familiar with these studies, not because I think they are better than studies conducted by other researchers.

In this volume, I focus on the technology of diary studies: how to design them; how to conduct them; and perhaps most important (and nearest and dearest to my heart), how to analyze the data produced in a diary study. I spend relatively more time on data analysis because I think it is the weakest link in the present chain of scientific inquiry. Although designing and conducting a diary style requires being sensitive to different aspects of a study than is the case for survey and experimental studies, designing and conducting diary studies has much in common with designing and conducting studies using other methods. For example, regardless of how data are collected, self-report items need to be written unambiguously, using language that is appropriate for the participant population.



In contrast, analyzing the data collected in a diary study requires methods other than the variants of ANOVA and regression that are typically used to analyze the data from survey and experimental studies. Although I discuss other options, most diary studies are probably best conceptualized as some type of multilevel data structure in which diary entries/records are treated as nested within persons. Unfortunately, many researchers are unfamiliar with how to conceptualize relationships within the multilevel context and are unfamiliar with conducting multilevel analyses. Given this, I thought a specific focus on how to analyze the data produced by diary studies was needed.

### Why diaries instead of other methods?

Before going further, I should probably describe what a diary is and what is meant by the term “diary method.” It appears that the English word “diary” is based on the Latin *diarium*, which referred to a daily allowance of food or pay. By the early seventeenth century, the word “diary” had taken on its present meaning of some type of daily record of events maintained and updated by an individual. Although, technically speaking, diaries are a type of daily record, diaries as discussed herein do not need to use the day as an organizing unit. For now, it will suffice to note that personality and social psychologists tend to use diaries with two different types of organizing units. In some studies, people keep a diary that is organized in terms of a type of event. Every time a certain type of event occurs, a diary record is created. In other studies, people keep a diary that is organized in terms of time. Every time a certain amount of time passes (which could, but does not have to be, a day), a diary record is created. I discuss these techniques in detail later.

What these techniques have in common is the fact that an individual provides a description of his or her life on some type of regular basis. In fact, diary research is sometimes described as a type of “intensive repeated measures” design because repeated measures are collected on the same person in what some think of as intensively. You will also see the term “experience sampling” because, in some studies, samples of people’s experiences are examined. This is often abbreviated ESM (Experience Sampling Method). Other terms you may see include “ecological momentary assessment” (EMA), a term that is used frequently in the health sciences. Each of these terms has been used to refer to different types of diaries.

Importantly, for the types of diary methods I discuss in this volume, data are collected “in vivo,” in the course of people’s everyday lives. The focus of most diary studies is not on the dramatic or the unusual aspects of a person’s life. Rather, diary studies tend to concern the natural ebb and flow of a person’s life, which invariably contains a mixture of the dramatic and the mundane. The value of collecting data in vivo is described in detail by Reis (2012). These advantages include the realism of the settings and contexts within which phenomena are

## INTRODUCTION

studied and the inclusion of the mundane and *apparently* unimportant aspects of life. As I discuss later, maintaining a diary should not interfere with a person's life. To the extent that it does, maintaining a diary destroys the very thing it is designed to study.

Fair enough, but why not study everyday life (or everyday behaviors) in the laboratory? For many social and personality psychologists, perhaps a bit more for social psychologists, the control and resulting clarity of inference provided by laboratory methods is very attractive. Independent variables can be manipulated, extraneous situational factors can be controlled or eliminated, and causal relationships can be isolated and decomposed. For many researchers, whatever loss of external validity occurs when using laboratory methods is more than offset by the gains in internal validity such methods provide.

Such a calculus has been, and remains, attractive. Nevertheless, by their very nature, laboratory methods are limited in important ways. Some phenomena simply cannot be studied in the laboratory because they cannot exist or be created within the controlled, artificial environment of the lab. Although we can induce feelings of sadness in people, we cannot make them depressed. Similarly, although we can create positive interpersonal bonds, we cannot make people fall in love. Anyone who has been depressed or in love can readily and easily recognize the qualitative differences between these *in vivo* experiences versus the fleeting states of mind that are created in most experiments.

Even if we could create such states of mind in controlled settings, some phenomena of interest to social and personality psychologists are destroyed when they are controlled. The individuals with whom someone affiliates and how often these people are seen cannot be studied in a controlled setting. Similarly, the types of situations people chose (or find themselves in) and how they react to these situational influences cannot be studied in the lab. Certainly, we can study such topics in the lab in a limited way. We can provide people the opportunity to select others with whom they will interact or to select situations and see how they respond, but it is difficult to imagine that the choices that can be made available in a lab setting can represent the range of options in people's real lives.

Okay, but: "Why do data need to be collected every day or a few times each day? Why not give people a single questionnaire asking them what happened during some period of time such as a week?" The answer is clear. The repeated measures collected in diary studies provide more accurate descriptions of everyday experience than those obtained from single assessment questionnaires that often ask respondents to recall events, thoughts, and feelings over lengthy, sometimes unspecified periods of time. Such vagueness is prone to all sorts of memory biases.

Although there are various lists of why such retrospection is bad (or, more charitably, is subject to error), I think the summary provided by Reis and Gable (2000) does a good job of capturing the important points. I summarize these below, and I add one more point that builds on these. The common theme to these specific

problems is perhaps best summarized by the Chinese proverb: “The faintest ink is more powerful than the strongest memory.” This proverb also reflects the considerable research demonstrating that memory (retrospection) is a constructive process. The eye is not a camera, and the ear is not a tape recorder. When we remember, we create memories as much, and perhaps more, than we recall memories. In terms of providing descriptions of what has occurred, the longer the time between the event and the description, the more extraneous factors (factors not related to the event itself) can influence the description.

1. *Recency* – Recent experiences are recalled better than temporally distant ones. Moreover, whatever occurs between an experience and the report of that experience is likely to influence the report. The longer the time between an experience and the report of that experience, the greater the opportunity there is for intervening events to influence the report. Such influences may be particularly important when individuals are describing separate experiences of the same type, such as a series of social interactions with the same person. An interaction I just had with a friend of mine may influence my recollection of the interactions I had with him during the past week.
2. *Salience* – More distinctive, important, or personally relevant experiences are recalled better than common or mundane experiences. In addition, more distinctive, important or personally relevant experiences are likely to influence reports of more mundane experiences. Such possibilities are particularly important when considering the relative influence of positive and negative experiences. Considerable research indicates that negative stimuli are more salient than positive stimuli. An argument with a friend or a bad day at work may be remembered better than a pleasant interaction of a normal day at work. In addition, important experiences are likely to influence reports of less important experiences, creating a sort of global report that may not reflect the nature of less important experiences or how people felt or thought during those experiences. An argument with a friend may influence my description of other, less dramatic interactions I had with him. An unusually bad day at work may lead me to think about other days differently.
3. *Sense-making* – Individuals have various implicit theories and they hold various stereotypes, and these implicit theories and stereotypes help people organize reality. As a result, individuals tend to reconstruct their recollections in terms of their general beliefs, often overriding reactions they would have provided at the time they experienced something. As the time between an experience and a report increases and as the number of reports increases, the distinctiveness of each experience decreases. As the distinctiveness of an experience per se decreases, the influence of implicit theories and stereotypes on reports of that experience is likely to increase.
4. *Present state of mind* – Individuals’ present or immediate state of mind such as moods or attitudes can influence memories, including reports of previous experiences. Such influences can take two forms. Present states of mind may influence the specific experiences that are recalled. For example, when in a good mood, people may be more likely to recall positive experiences, whereas when they are in

## INTRODUCTION

a bad mood they may be more likely to recall negative experiences. Present states of mind may also influence reports of experiences per se. For someone who holds strong negative attitudes about a certain ethnic group, a retrospective report about an interethnic interaction may be more negative than a report that is provided more contemporaneously.

- 5 *Making distinctions* – As discussed by Nezlek and Schaafsma (2010), compared to reports generated in diary studies, retrospective reports tend to reflect more strongly the operation of a single (often, hedonic) dimension. When thinking about multiple experiences in the perhaps distant past, people are less likely to make distinctions among different aspects of those experiences compared to the distinctions they make when providing more immediate reports. The hedonic (good–bad) dimension seems to be the one that is the most salient across time. The fact that a retrospective report asks people to distinguish different aspects of an experience (e.g., satisfaction with an interaction vs. how influential an individual felt during an interaction) does not mean that people make such distinctions when providing retrospective reports. For the reasons just discussed, when providing retrospective reports, people do not tend to distinguish different aspects of an experience as well as when they are making a contemporaneous report. For example, relationships (e.g., correlations) between different evaluations tend to be stronger for retrospective reports than they are for contemporaneous reports.

All of this is not to say that single assessment, retrospective measures are not valuable. If I ask you how your life was last week and you reply “Miserable,” that is meaningful and potentially valuable in terms of understanding your life. On the other hand, it may not be a good measure of how you felt each day of last week. On six of those seven days you may have felt fine, but on the last day of the week something terrible happened, which colored your impression of the other days. Perhaps more important, it may also be that some type of measure based on individual reports from all seven days provides a more useful measure (in terms of relationships with other criteria) than the single assessment, global retrospective report.

Although both laboratory-based studies and single assessment surveys are limited in important ways, I have no desire to denigrate or criticize these methods per se. They are valid and informative ways of understanding the human condition. In this volume, I will be describing how to design and conduct diary studies and how to analyze the data they produce without constantly comparing diary and other methods *mano-a-mano*.

### Levels of analysis

In addition to providing more ecologically valid and accurate descriptions of people’s lives than lab and survey methods, diary methods also provide a better basis for examining psychological phenomena at different levels of analysis.

When describing diary studies, the phrase “levels of analysis” refers to two mutually defining aspects of the data collected in a study, the technical and the conceptual. Moreover, for most diary studies, there are two levels of analysis, the person level and the diary level. There can be more than two levels, and I discuss such possibilities later.

Technically speaking, level of analysis refers to the focus of the measures that are collected in a study. Person-level measures include those that describe something about a person that is presumed not to vary over the course of a study, such as an individual’s sex and personality traits. In contrast, diary-level measures include those that describe something about a diary entry (or record – the target of the diary *per se*), and diary level measures are presumed to vary across the diary entries for individuals. A measure of mood that is collected every day is a prototypical diary-level measure. I discuss person- and diary-level measures later in separate sections.

Conceptually speaking, level of analysis refers to the level at which a phenomenon (including relationships between constructs) is thought to reside. This also includes the type or nature of the processes that relationships between measures are meant to represent. Relationships at the person level involve person-level measures, and they concern person-level processes. For example, the relationship between trait self-esteem and trait anxiety indicates if people who are more anxious tend to have a lower sense of self-worth compared to individuals who are less anxious. In contrast, relationships at the diary level involve diary-level measures, and they concern diary-level or within-person processes. For example, the relationship between a daily measure of mood and a daily measure of stress indicates if people’s mood is poorer on days when they experience more stress than on days when they experience less stress.

It is critical to keep in mind that relationships at these two levels of analysis are both technically and conceptually distinct. Technically, as I explain later, relationships between two constructs measured at the person level are mathematically unrelated to relationships between these same two constructs measured at the diary (or within-person) level (Nezlek, 2001). Perhaps more important, conceptually, relationships at the two levels of analysis may represent different psychological phenomena (Affleck, Zautra, Tennen, & Armeli, 1999).

In diary studies, multiple observations are collected for each person, providing a basis for analyzing within-person relationships (diary-level relationships). Although it is theoretically possible, in the lab, it is difficult to collect the data that are needed to provide good estimates of within-person relationships such as that between stress and anxiety. How many different experimental conditions can people experience before the quality of the data they provide becomes questionable? Although such repeated measures could be asked in a survey, how well can people provide retrospective accounts of events and internal states for a sufficient number of occasions to provide a basis to estimate within-person relationships?

Whatever shortcomings diary studies have, it seems that they provide the best opportunity to examine within-person processes. Moreover, they can do so in a way that maximizes the ecological validity of the study. Truly, they can study “life as it is lived” (Bolger, Davis, & Rafaeli, 2003).

Note that, in the literature, person level is sometimes referred to as “between person” to provide a more fitting complement to “within person.” Throughout this volume, I will use person level and between person interchangeably, and I will use diary level and within person interchangeably.

### Does keeping a diary per se influence what a person reports?

It appears that the multiple real-world observations collected in diary studies can provide unique advantages compared to other methods. If so, we should consider the potential disadvantages of asking people to answer the same questions on a repeated basis. One possible disadvantage is that providing responses in such a fashion might influence what people report. Such an influence could concern responses per se. Do people tend to answer questions differently at the end (or middle) of a diary than they do at the beginning? Do they become sensitized to certain events and conditions and increase the number of events they describe over time? Or, do they become desensitized, bored, or tired, and describe fewer events in less detail over time? Alternatively, and perhaps more fundamental, does keeping a diary change a person’s life, aside from what they report about it? For example, if I record all my social interactions, do I change with whom I interact or how I think about my interactions as I see as an objective description of my social life developing before my eyes?

For the most part, researchers have assumed that diaries are non-reactive in both senses of the term as described above, and there is some research supporting such an assumption. Although self-reports of reactivity do not constitute a case on their own, in post-study interviews with participants in studies using the Rochester Interaction Record, of which I have been part, participants have rarely reported any sort of meaningful change as a result of maintaining the record. Some have reported an increased awareness of what they were doing: for example, “I never realized how much time I spent with the same group of people.” Such individuals did not report, however, that they changed their behavior as a result of this awareness. Similarly, I have compared the first and second halves of the data I have collected in some of my daily diary studies (usually a week for each half). Although I have found differences between the halves in means of some measures, these differences were quite small, and there were no differences in within-person covariances between measures, and such covariances were the focus of my hypotheses.

Moreover, in a series of studies discussed by Hufford (2007), researchers did not find any evidence for reactivity. Admittedly, these studies did not concern all types of diaries focusing on a wide range of topics. They primarily concerned momentary assessment studies of patient-reported outcomes such as pain. Nevertheless, to my knowledge, no one has demonstrated that maintaining (per se) the types of diaries that are discussed in this book meaningfully changes the behaviors of the diary keeper. For a more thorough discussion of this topic (that reaches a similar conclusion) see Barta, Tennen, and Litt (2012). Barta et al. also suggest, however, that not enough is known about reactivity because it has not been studied that carefully and that more research is needed before we can safely assume that maintaining a diary does not lead to systematic changes in the behaviors of the diary keepers.

Nevertheless, some research suggests that under some circumstances, maintaining a diary may lead to changes in those who are maintaining it. As might be expected, much of this research has been conducted by researchers with a clinical interest or focus; however, the issues this research raises and addresses may be relevant to the interests of some social and personality psychologists. As discussed by Carter, Day, Cinciripini, and Wetter (2007), maintaining a diary “increases the patient’s awareness of the frequency, patterns, and circumstances attendant to a target behavior” (p. 293). They further discussed how the feedback provided by a diary may reduce undesirable behaviors and increase desirable behaviors. Note that they discuss “patients,” and mention “target behaviors.”

What is important to note about the type of research discussed by Carter et al. is that invariably the diary keepers in such studies had goals of some kind. They were patients, under care, and they (presumably) wanted to make some type of change in their behavior or they wanted to maintain some type of change they had made. There were explicitly desirable behaviors such as exercise or compliance with medication protocols, and there were explicitly undesirable behaviors such as smoking or eating too much fat.

In contrast, in many (or perhaps most, or virtually all) diary studies conducted by personality and social psychologists there are no explicitly desirable and undesirable behaviors. Quite the opposite. Researchers typically make it clear to participants that there are no desirable or undesirable behaviors. Participants are told that the researchers are simply interested in their daily life per se. In fact, to the extent possible without creating issues about reactivity, participants are often reminded that they should not change their behaviors or routines because they are maintaining the diary.

So, the available evidence appears to suggest that when people maintain a diary without an explicit goal or end state in mind, maintaining a diary does not seem to have a meaningful influence on their thoughts, feelings, or behaviors. In contrast, diaries can be considered as a type of intervention when there is some type of desired goal or end state in the mind of the participant.

## Additional resources

Regardless of your level of familiarity with diary research, you may find the following sources helpful. This list is not meant to be complete, although the articles and chapters do cover most of the major issues. Moreover, for readers who are not familiar with diary-style research or certain aspects of such research, these are “good places to start.” I have omitted some earlier citations whose content or contributions were either dated or covered in more recent citations. Later, I present a similar list of sources for data analysis.

- Wheeler and Reis (1991) – A discussion of the relative advantages and disadvantages of different data collection protocols.
- Affleck, Zautra, Tennen, and Armeli (1999) – A discussion/explanation of why it is important to distinguish relationships at different levels of analysis, specifically between person vs. within person.
- Gable and Reis (1999) – A review of using within-person designs (mainly diary methods) to study personal relationships.
- Reis and Gable (2000) – A review of methods of studying everyday experience.
- Bolger, Davis, and Rafaeli (2003) – A review of various aspects of diary research with a focus on methods and analytic techniques.
- Stone, Shiffman, Atienza, and Nebeling (2007a) – A review of the development of ecological momentary assessment (EMA) techniques in the health sciences, including a rationale for them.
- Nezlek and Schaafsma (2010) – A discussion of the advantages of using diary style methods to study interethnic contact.
- Reis (2012) – A broad overview and rationale for using diary-style methods.

In addition, a forthcoming handbook, edited by Mehl and Conner (2012), has chapters concerning various aspects of diary-style research. For those interested in EMA with a focus on applications in health I recommend Stone, Shiffman, Atienza, and Nebeling (2007b), an edited volume based on a 2003 meeting convened by the US National Cancer Institute.

The important advantage of the present volume over other sources (at least for personality and social psychologists) is that I discuss aspects of diary research in terms that are particularly relevant to personality and social psychologists and do so in greater depth and detail than is available elsewhere (with all modesty). Moreover, in this volume, discussions of design, analysis, and reporting issues are fully integrated. Finally, and consistent with the focus of the series of which this volume is a part, I discuss topics in a fashion that makes them accessible to the non-expert.