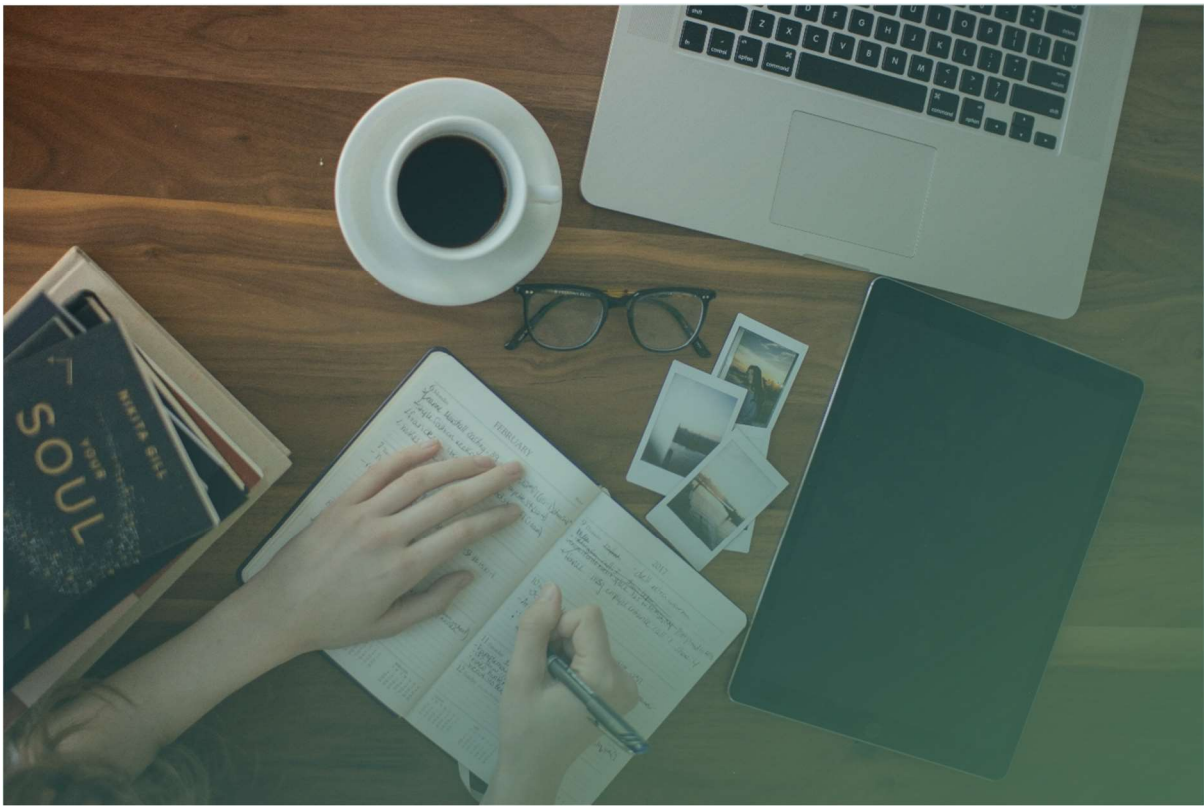
Lades, L. and Kelly, A., 2020

EnvEcon
Decision Support



Using the Day Reconstruction Method to Inform Policy

A practical guide

EPA RESEARCH PROGRAMME 2014–2020

The EPA's current Research Programme 2014–2020 is built around three pillars - Sustainability, Climate and Water.

More information about the EPA Research Programme can be found by visiting the EPA website where you can subscribe to the Research Newsletter. This provides news and updates about research calls, events and publications that are of relevance to researchers and other interested parties.

You can also follow EPA Research on Twitter @EPAResearchNews for the information and developments about the Research Programme and its projects."

ACKNOWLEDGEMENTS

This project is funded under the EPA Research Programme 2014-2020. The EPA Research Programme is a Government of Ireland initiative funded by the Department of Communications, Climate Action and Environment. It is administered by the Environmental Protection Agency, which has the statutory function of co-ordinating and promoting environmental research.

Dr Kelly's engagement in this work is also supported by the DCCAE.

REFERENCE

Lades, L. and Kelly, A. (2020), DRM - Informing policy-making with data from everyday life:

A practical guide. Dublin. EnvEcon

TABLE OF CONTENTS

EXECUTIVE SUMMARY

This guide aims to facilitate the application of naturalistic monitoring to the understanding and development of better behavioural policy interventions. Naturalistic monitoring is a technique that generates data from everyday life. By observing people's behaviours and experiences "in the wild". We focus on the Day Reconstruction Method (DRM) as one particular naturalistic monitoring tool that can be used cost-effectively at scale.

The guide can be of value to you if you are looking into analysing everyday experiences, behaviours, or decisions occurring in real-world settings and if you are interested in what actually happens, rather than in what can happen under specific circumstances. To the contrary, if you are interested in stable aspects of life such as attitudes or general preferences, naturalistic monitoring should not be the measurement tool of your choice.

The guide summarises the key characteristics of naturalistic monitoring in section 2. It presents and discusses the Day Reconstruction Method in section 3. It describes six key steps to be taken when designing Day Reconstruction studies in section 4. This is the core of the guide. The key steps are:

Step 1: Making sure that the DRM is the right method.

Step 2: Defining the relevant sample and time frame.

Step 3: Designing the DRM diary.

Step 4: Linking the diary with the follow-up questions.

Step 5: Design the follow-up questions.

Step 6: Analysing the DRM results.

Finally, the guide demonstrates how the Day Reconstruction Method can be used to evaluate and design policy interventions showing some example results in section 5 and concludes in section 6.

1 INTRODUCTION

Policy often seeks to change the behaviours of people in their everyday lives. Some policies try to make people eat healthy foods, exercise more, use less fossil fuels, recycle their rubbish and so on. The introduction of policy interventions is often paired with efforts to measure subsequent outcomes (e.g., changes in travel patterns or reductions in energy use). However, the measurement of the changed outcomes contributes little to the understanding, refinement, and application of effective policy interventions on different scales, at different times, and in different contexts.1,2 In short, knowledge about outcome changes following policies does not necessarily explain how the changes came about. Designing and deploying better policy, however, requires an understanding of causal and coincidental mechanisms. To understand the mechanisms by which policies influence everyday behaviour, data from everyday life is needed.

Naturalistic monitoring is an established research technique that generates such data from everyday life.[3] It is built upon the observation of people's behaviours and experiences "in the wild", and captures situation-specific thoughts, feelings, and behaviours as they arise in daily life. A common application of naturalistic monitoring is to identify the everyday activities during which people are most and least happy.[4] However, naturalistic monitoring can also be used to measure decisions and the situation-specific variables that influence these decisions.[5,6] This can help us to better understand how decisions are influenced by situational context-factors that can be modified by policy.

Naturalistic monitoring has been enthusiastically implemented in psychological research and happiness studies. However, despite the great potential, the tool is not yet widely used to measure the effects of policies on everyday behaviours and experiences. This brief's main objective is to facilitate the application of naturalistic monitoring, specifically the Day Reconstruction Method (DRM), to the understanding and development of better behavioural policy interventions. The guide is structured as follows:

- Section 2 summarises the key characteristics of naturalistic monitoring.
- Section 3 presents and discusses the Day Reconstruction Method.
- Section 4 defines the six steps to be taken when using the Day Reconstruction Method.
- Section 5 demonstrates how the method can be used to evaluate and design policy interventions.
- Section 6 concludes.

If you are interested in collaborating on a DRM study, please contact the authors at admin@EnvEcon.eu.

2 WHAT IS NATURALISTIC MONITORING?

Naturalistic monitoring describes the observation of people's behaviours and experiences in their natural environments, i.e. where most of decision-making takes place. It includes self-reported measures but also more objective observations, for example of people's psychobiology such as their heart rate or other data that can be captured from smartphones or other portable devices.[7] Naturalistic monitoring studies have three core characteristics:[3]

1. Naturalistic monitoring studies provide **data from people's every-day lives** and provide information about what actually happens. Other study-techniques, such as lab experiments and many surveys, put individuals in situations they would not encounter in their normal lives. Naturalistic monitoring tells us what actually happens, while many lab experiments and surveys deal with what can happen.
2. Naturalistic monitoring studies are **conducted in (or close to) real-time** and hence provide data on situations as they are experienced in the moment rather than how people recall them in distant memory. This can improve data quality as it overcomes memory biases that can influence how people respond to survey questions. Naturalistic monitoring studies aim to activate episodic memory (that is activated by questions such as "Did you enjoy your last meal?") rather than semantic memory (that is activated by questions such as "What is the capital of France?").
3. Naturalistic monitoring studies **measure concepts repeatedly** and hence allow analysis of intra-individual changes. While many experimental and survey studies focus on differences across individuals, naturalistic monitoring can help with understanding how the same individual differs across situations. This allows us to analyse how the same individual interacts with different decision-making contexts that are potentially affected by policy interventions.

3 THE DAY RECONSTRUCTION METHOD (DRM)

3.1 What is the DRM?

The Day Reconstruction Method (DRM) was developed as a cost-effective means of naturalistic monitoring by Nobel prize winner Daniel Kahneman and colleagues. The goal was to establish a methodology to assess how people spend their time and how they feel in their everyday lives that can be employed on scale in large populations. The DRM is widely used to measure subjective well-being in everyday life and has been used both in large representative population surveys,[8,9] and smaller scale studies focused on specific topics.[5,10–14]

The DRM is comprised of a two-step survey:

- **Step 1:** In the first step, participants reconstruct their previous day ("Yesterday") in a diary. In the diary, participants are asked to divide their day into "episodes", as if each episode was a scene in a movie. Participants give each episode a name, such as "Driving to work" or "Having lunch" and are asked to reflect upon what they did and how they felt during each episode. Adding a diary before the follow-up questions is essential as it minimizes memory biases. For example, it is difficult to give an accurate answer to the question "How did you feel yesterday at 3.30pm?". However, after having completed a diary of yesterday, responding to the same question is much easier and answers are more accurate.
- **Step 2:** In the second step, participants complete a structured survey in which they are asked follow-up questions about each episode. These follow-up questions can deal with any themes the researchers wish to assess.

3.2 What are the advantages and disadvantages of the DRM?

Advantages

The DRM has several characteristics that make it a valuable tool for use in policy and business-relevant research. Amongst these characteristics are:

- **The DRM provides data from the real world.** Understanding the real world is arguably more important than understanding behaviour in the lab or other artificial environments. This is an advantage as it is not certain whether lab or survey findings generalise to the real world. For example, testing for the influence of experiences (such as stress) in the lab might differ qualitatively from the same influence in the real world. Moreover, some studies are difficult (or unethical) to conduct in the lab. For example, it is ethically unacceptable to artificially create disagreements with one's co-worker or encourage excessive behaviours such as binge eating or drinking in the lab. If one is interested in these situations, measuring them as they occur naturally is often the only possibility.
- **The DRM is cost-effective.** DRM studies are comparably easy to conduct and can be managed by most researchers without the help of technicians. Moreover, the participant burden is relatively low as DRM studies can be completed in a single study session.
- **DRM studies can be scaled.** It is possible to conduct DRM studies with large, nationally-representative samples. As such, sub-populations can be identified for which policies might be more or less effective.
- **The DRM does not interfere with people's lives.** Smartphone-based study designs ask study participants to stop with whatever they are doing and to answer questions. In some domains (e.g. driving, teaching, or working) people cannot make breaks to respond on the smartphone. DRM studies do not suffer from this high interference.
- **DRM data can be for the entire day.** This allows to obtain more than just snapshots of daily life, and provides information about, for example, how people spent their entire 24 hours.
- **DRM data has high validity.** Data gathered from the DRM is often systematically similar to the data obtained through smartphone studies. Even when the recalled DRM data diverges from the real-time smartphone data, the DRM data might be more important because many decisions people make are informed by the memory of experiences rather than how they objectively felt the experiences.

Disadvantages

The DRM also has limitations, and researchers and policy-makers need to be aware of these before designing research studies and evaluating DRM data for policy-purposes. The main limitations are:

- **DRM data is subjective and self-reported.** Whenever study participants respond to survey questions, a number of factors can reduce data quality, including dishonest reporting, social desirability bias, norms, self-image considerations, sensitivity of measurement instruments, and

reactivity to assessment procedures. There is a large literature on these issues that are present is all types if survey research and also in DRM research.

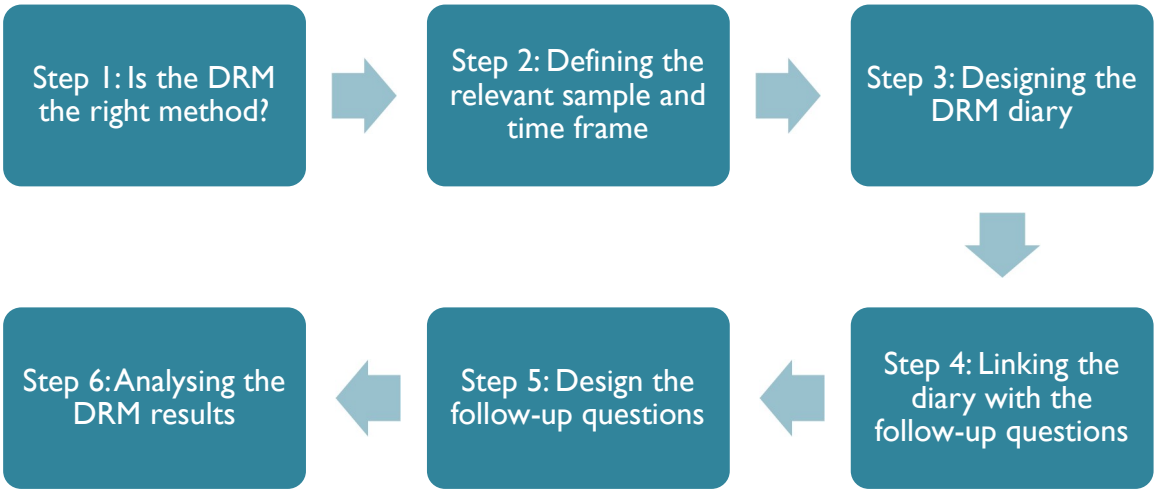
- **DRM date relies on the participants' memory.** While the two-step procedure of the DRM, with the diary as the first step, is designed to minimise memory bias, it is likely that some memory biases still exist in DRM data. Participants might rely on routines to infer yesterday's likely activities, they might use the peak-end rule to evaluate past experiences, and overstate previous emotions.15 However, these memory biases seem to be rather small as previous comparisons of experience sampling and day reconstruction data suggest that both methods provide similar data.16–18
- **The DRM is a relatively novel method.** As with every novel tool, there are unknowns and areas for improvement. For example, as will become obvious throughout this document, there are many variants of the DRM and methodological research has not yet identified the optimal design of DRM studies.

While we make a general case for informing policy-making with naturalistic monitoring techniques, we view the DRM as the most suitable method to be implemented by policy mainly due to its cost-effectiveness and scalability, especially when analysing large samples. We note, however, that the method has only been recently developed and more research on validity and best practices would be beneficial.

4 HOW TO DESIGN A DAY RECONSTRUCTION STUDY

This policy-brief main aim is to facilitate the use of the DRM to evaluate policies that aim to change every day behaviours. This section presents an accessible yet detailed guide to the use of the DRM. Specifically, we present the six key steps that need to be taken when using the DRM.

Figure 1. The six key steps when designing a DRM study



4.1 Step 1: Making sure that the DRM is the right method

The first step when designing a DRM study is to make sure that the DRM is a valid tool for the research question. There are cases when naturalistic monitoring is not the best option to answer a research question. As an example, naturalistic monitoring is not the best method to analyse why people prefer a certain car/mortgage/pension fund over another. These big, one-off decisions do not happen very frequently, are not characteristic of daily life, and the associated deliberation processes can be better understood using lab experiments and surveys. However, naturalistic monitoring is a very good choice when investigating behaviours such as smoking, eating, taking the car rather than the bus, or watching another episode of one's favourite TV-show. Such high-frequency decisions are often made on a daily basis intuitively with little effort and deliberation.

The DRM is the right tool:

- If the research question deals with everyday experiences, behaviours, or decisions occurring in real-world settings without much deliberation (e.g., when asking "How happy are people when commuting?").

- If researchers are interested in what *actually happens*, rather than in what *can* happen under specific circumstances (e.g., when asking "How prevalent is behaviour X in everyday life?").
- If the research investigates contemporaneous determinants of decisions (e.g., when asking "How does the music affect consumption in the supermarket?").
- If the research deals with the dynamics of life as it is lived, day-to-day, hour by hour (e.g., when asking "Are self-control failures more common in the evening compared to the morning?").

The DRM is not necessarily the right tool:

- If the researchers are interested in rather stable aspects of life where the within-person variability of multiple measures is small. When attitudes or general preferences are measured, global surveys should be used (e.g., when asking "Does the personality trait conscientiousness correlate with an environmentally-friendly lifestyle?").
- If researchers are interested in behaviours that follow from a deliberative decision-making process leading to rather stable answers. In these cases, surveys and lab experiments are adequate tools (e.g., when asking "What are the factors that explain purchases of electric vehicles?").
- If the researchers are interested in potential behaviours, i.e. things that can happen, and their causal determinants. In these cases, lab and field experiments are the method of choice (e.g., when asking "Does providing information about the energy-efficiency of a product increase the uptake of that product?").

4.2 Step 2: Defining the relevant sample and time frame

Composition of sample

After having identified the DRM (or another naturalistic monitoring tool) as a valid method to answer the research question, the researcher needs to identify the relevant sample. As in all survey-based research, a specific sub-sample of the population might need to be selected. If the research question, for example, relates to the current quality of public transport, there is not much use in selecting as sample of people who do not use public transport. If the research deals with the effects

of medication, participants who take the respective medication need to be approached (as well as other participants who are not under the same medication, but are otherwise similar and will be part of the control group).

Sample size

Another important decision concerns sample size, i.e. the number of participants who take part in the study. It is often a good idea to look at previous similar research and to aim for samples not smaller than in these previous studies. In other cases, nationally representative samples are needed and these studies will require hundreds and sometimes thousands of participants. It is possible to run DRM studies with such high numbers. However, the researcher needs to consider shorter versions of the DRM that have a lower respondent burden. When the research aims to test a specific hypothesis, a power analysis needs to be conducted in order to identify the minimum sample needed to be able to reject the hypothesis with sufficient certainty. There are a number of resources available on power analysis for multi-level models in naturalistic monitoring research.[19,20]

Which day was "Yesterday"?

DRM studies typically provide data about "yesterday". Hence, the researcher needs to make sure that the behaviour of interest actually happened yesterday. For example, if the research question relates to feelings during commutes to or from work, the study should not be conducted on a Sunday. If relevant behaviours are relatively infrequent, even more care has to be taken to select the right study days. For example, if the research is about consumer behaviour in supermarkets, it is important to select participants who went grocery shopping yesterday. In these cases, pre-screening surveys are useful. It is also possible to conduct Event Reconstruction Studies to analyse events that happened in the recent past (and maybe more than one day before).[21] For these studies it is essential to invite only participants who had the event under investigation in the recent past. Generally, it is possible to conduct the survey just once or multiple times with varying delays in between surveys.

4.3 Step 3: Designing the DRM diary

Recall that the DRM is a 2-step procedure. Participants make a number of diary entries in step 1 and then use these diary entries as a memory aid in step 2 where they answer follow-up questions. The key aim of the diary in step 1 is to maximize the extent to which participants re-live their previous day. This process is called re-instantiation and it guides the respondents to recall yesterday's experiences, behaviours, and decisions. It encourages participants to respond with their actual experience in mind. Without re-living the past experience, it is likely that respondents draw on intuitions shared by the researchers so that findings appear congruent and logical. If the re-instantiation is not successful, respondents might also use the present as a benchmark to reconstruct past behaviour, experience, and decisions. It is therefore important to allow participants some time to complete the diary so as to facilitate a successful re-living of yesterday's experiences. Figure 2 shows an example text that illustrates one way to communicate the 2-step procedure to study participants.

Figure 2: Example for an DRM introduction text

Helping you to remember

We are interested in what happened **yesterday** (Wednesday, September 6th) around **19:00h** and some time thereafter. The time shown is based on a 24 hour clock.

As people find it difficult to remember what exactly they did and felt a day before, we like to proceed in 2 steps:

- Firstly, you will complete a short diary (see below).
- Secondly, you will answer some follow-up questions (we will show you your notes from step 1 again).

Diary

To complete the diary in the boxes below, please think about **what you did** and **how you felt** yesterday around **19:00h**. Think of this as an **episode in a movie**. Such an episode typically lasts between 15 minutes and 2 hours. The end of an episode might be going to a different location, or ending one activity and starting another.

After the first episode, repeat the same for the **directly following** episodes 2 and 3.

What you write in the boxes only has to make sense to you. We will see what you wrote, but we will not analyse it.

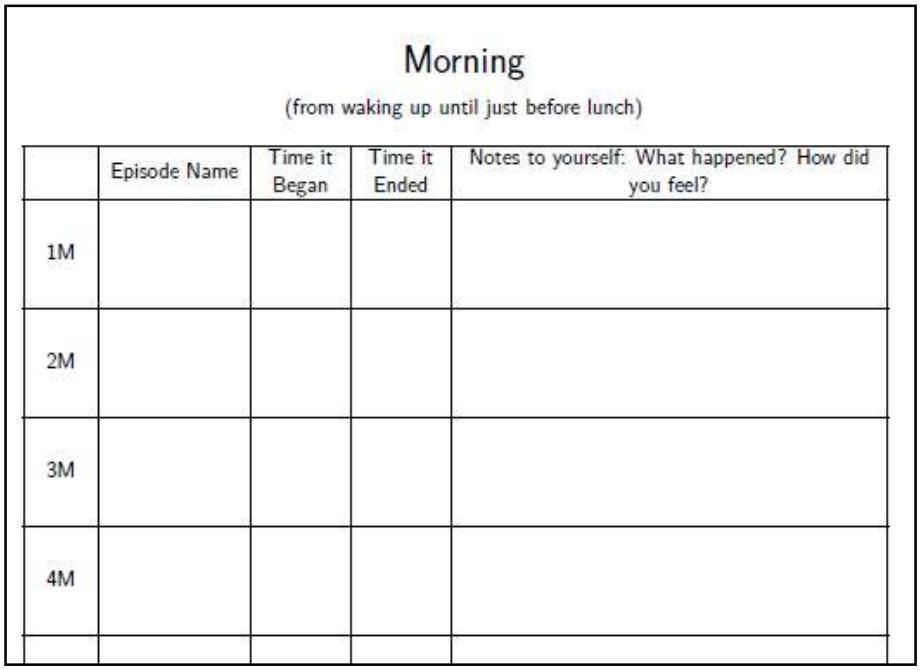
You have **50 characters max** for each box.

Feel free to have a look at your calendar or mobile phone to recall what happened yesterday around **19:00h** and in the two episodes afterwards.

Mode of administration

The diary can be administered in various ways, including on paper, online in a web-browser, or on a mobile phone. An advantage of asking people to fill out the diary on paper is that the diary can be private, for the participants' eyes only. Researchers can ask the participants to take their diary home after the study, which increases the chances that participants reconstruct their day without being concerned about the impressions they make on the researchers. In some cases, it is not possible to conduct a paper-based diary as many DRM studies are completed online. In these cases, the researchers can inform the participants that the diary notes will be seen by the researchers, but the researchers will not analyse them. Figure 3 present an example diary that can be handed out to the participants on a piece of paper.

Figure 3: Example for a paper diary for the morning. Similar sheets should be made available for the afternoon and the evening. Each part of the day allows for 20 entries



Morning

(from waking up until just before lunch)

	Episode Name	Time it Began	Time it Ended	Notes to yourself: What happened? How did you feel?
1M				
2M				
3M				
4M				

Length

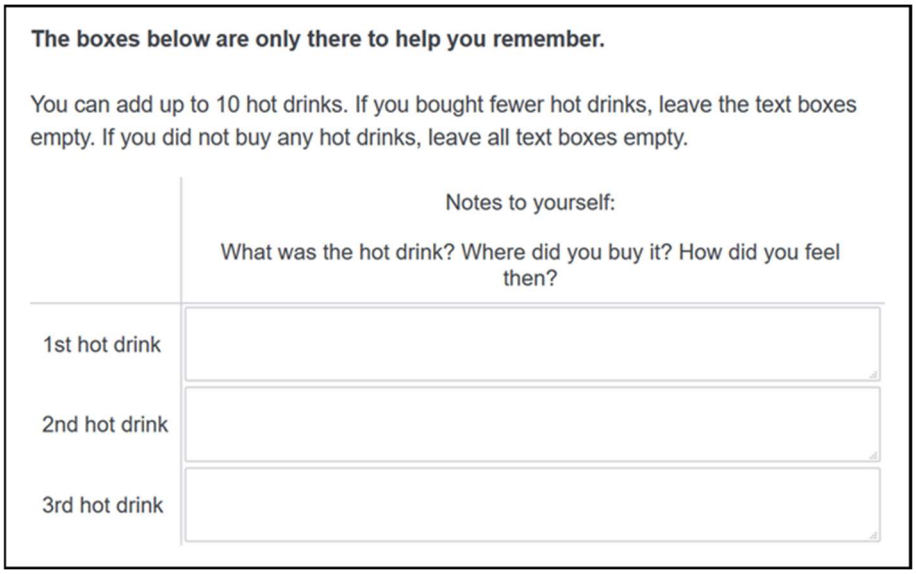
The optimal DRM study asks participants to reconstruct their full day. However, writing a diary for the full day can take a significant amount of time and effort. As reducing participant burden is often a good idea to minimise attrition and maximise sample size, some studies benefit from shorter diaries. It is possible, for example, to ask respondents to only recall yesterday's morning, afternoon, or evening as illustrated in Figure 3. Alternatively, participants can be asked to recall only a minimum of 5 episodes starting at a random point of the day as illustrated in Figure 2. Note that the minimum of 5 observations per individual is suggested in order to estimate a linear model with an intercept, a slope for time, a slope for a prior value of the outcome, and a slope for one putative causal antecedent.[19]

Partitioning of the day

In most DRM studies, participants are asked to partition their days individually into various episodes. They are often told that the typical length of an episode is between 15 minutes and 2 hours. However, it is also possible to pre-define certain time intervals, for example 15-minute

intervals, and ask participants follow-up questions regarding each of these intervals. The time intervals can be shortened or lengthened depending on the research question. For example, if the research is about rather short-lived aspects of everyday life, such as specific decisions and their psychological and situational correlates, many shorter episodes are preferable over fewer longer episodes. The pre-defined time intervals also reduce individual heterogeneity in terms of the number of reported episodes.[15] It is also possible to deviate from the episode as the unit of analysis according to which the second phase of the DRM is structured. For example, one can ask participants to list all activities (not episodes) they engaged in yesterday and follow up with questions for each activity. Finally, it is possible to ask participants to recall specific activities (such as the purchase of hot drinks as in Figure 4).

Figure 4: An example of the diary based on the hot drinks that participants had purchased yesterday.



4.4 Step 4: Linking the diary with the follow-up questions

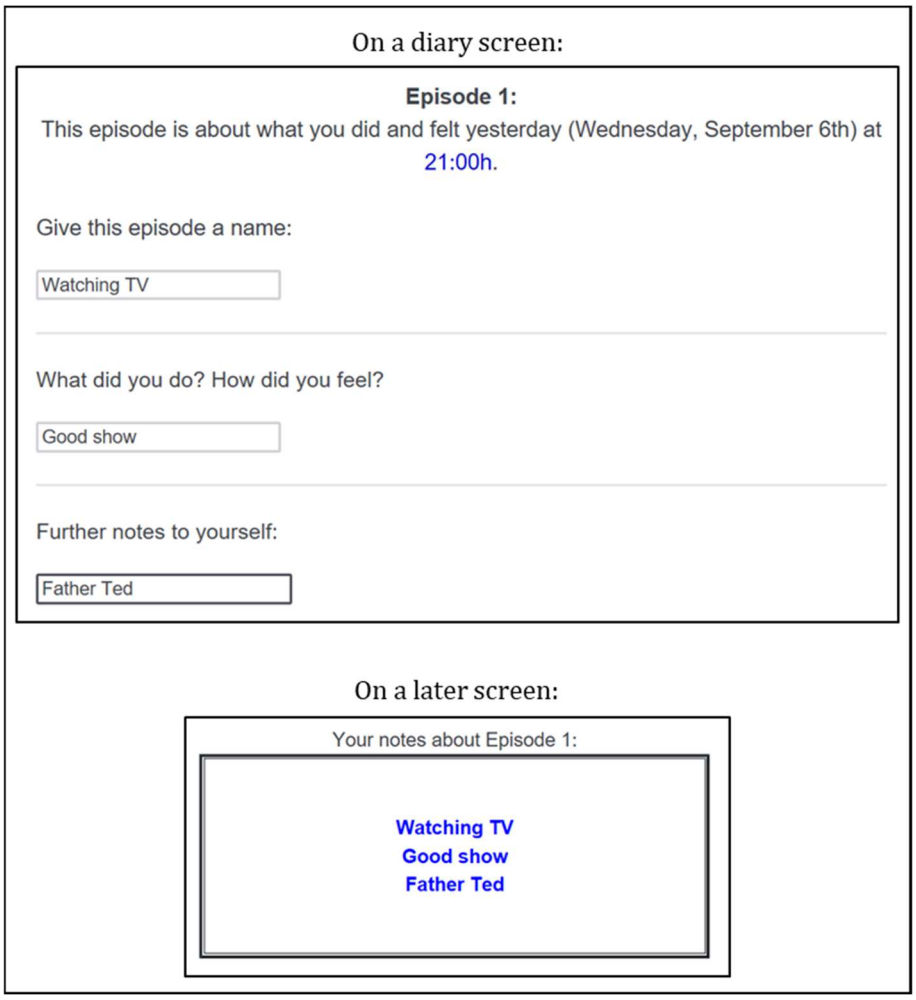
A key challenge in DRM research is to make sure that participants use their diary from step 1 when answering the follow-up questions in step 2. Only if the diary and the follow-up questions are linked in a meaningful way, the re-instantiation of yesterday will affect the answers participants give.

Mode of administration

Linking diary and follow-up questions differs by administration mode. If both the diary and the follow-up questions are answered on pieces of paper as in the seminal study by Kahneman and colleagues (2004), researchers can ask participants to refer to the pieces of paper participants used when completing the diary. In the study by Kahneman and colleagues, participants received one sheet of paper with follow-up questions for each episode. Since participants entered 15 episodes on average for each day to the diary, a sufficient number of sheets of paper need to be provided. For each episode, the researchers asked participants to take a look at the respective diary entry and to indicate a code (such as 1M, where the 1 stands for the first episode and the M for morning) at the top of the sheet before answering the follow-up questions. This procedure made sure that participants answered the follow-up questions with the respective episode on mind.

A similar system can be used when the diary is completed on paper and the follow-up questions are answered online. We recommend this procedure as the answers do not need to be transcribed before analysing the data. Moreover, survey software can loop questions several times, which is convenient as the follow-up questions only need to be changed once. If both the diary and the follow-up questions are completed online, participants should see their diary notes on every screen on which they respond to follow-up questions. Figure 5 presents an example where the diary entries are shown to the participants just before they answer the follow-up questions. Additionally, it can be helpful to ask participants to answer an episode-specific question (e.g., “Give this episode a name”) at the beginning of the page with the follow-up questions in order to make sure participants think about this specific episode when responding.

Figure 5: Example for a link between the diary and the follow-up questions.



Selection of episodes

It is possible to ask follow-up questions for all episodes described in the diary, or only for a subset of these episodes. One can randomly select some episode, pick episodes from specific points in time, or even ask respondents to choose the most important episode from yesterday. Depending on the research question, researchers can zoom in to explore single episodes in more detail or zoom out and explore all episodes in less detail. If the diary is partitioned according to different activities, it is also possible to ask follow-up question only to specific activities (e.g., commuting, working, or having social interactions).

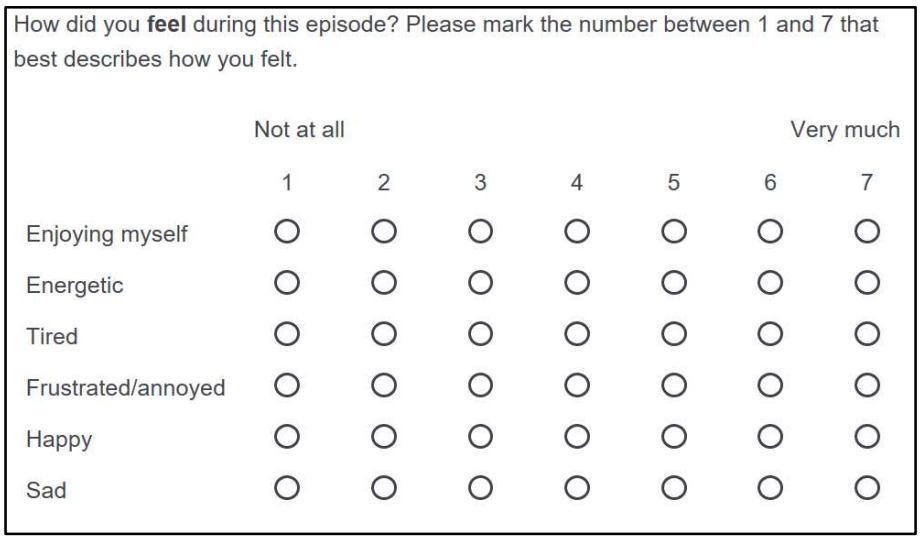
4.5 Step 5: Design the follow-up questions

A key element of every DRM study is to design appropriate follow-up questions to ask in step 2 of the study. While the same rules apply that are relevant in any kind of survey research, there are some considerations for question-design that are specific to naturalistic monitoring research.

Design the questions

Maybe the most important rule when designing questions for DRM studies is to keep the questions as short as possible. Since participants answer the same question multiple times, long questions can become burdensome and annoying to answer. It is also a good idea to use dynamic questions that go into more detail, but only if a condition was triggered by a previous question. Care should be taken to design the question that askes for the time of the day, as time of the day is an important variable in the analysis of DRM data (e.g., making sure participants give a clear indication whether they are describing an episode taking place at 9am or 9pm). Figure 6 shows a question that is maybe the most often asked question in DRM studies.

Figure 6: Example of typical momentary experience questions asked for each episode.



Converting questions

Many survey questionnaires used in the behavioural and social sciences are validated using psychometric methods. However, these validations usually focus on how individuals differ from each other and assume that these differences are rather stable. A key benefit of DRM research is that the effects of real-life contexts on decisions can be investigated; such contexts can result in very different patterns within the same individual. One strategy to design survey questions for use in DRM studies is to transform validated global survey questions into questions that are meaningful in the context of everyday life. In order make this transformation, researchers can identify global survey questions that are relevant to their research question, selects the survey items that are top-loading, and change the wording from global (e.g., "How important is recycling to you?") to momentary (e.g. "How did you get rid of the coffee cup?"). Afterwards, the new questions should be cognitively tested and piloted to make sure the questions are reliable and that participants interpret them correctly.

4.6 Step 6: Analysing the DRM results

In DRM studies, individuals provide multiple measures of the same construct over time. This is sometimes called "intensive longitudinal designs". Having data with repeated observations over time nested within individuals allows us to structure the data in at least three ways: (i) one row per individual, (ii) one row per episode, or (iii) one row per time unit. Different research questions can be answered with different data structures. Most statistical programmes have commands that make it rather straightforward to reshape the data from one data structure to another.

One row per individual

The data can be structured on the individual level so that each row represents one individual. In this case, each variable would appear multiple times in the same row. For example, if participants recorded happiness ratings for up to 5 episodes, the happiness variable appears in 5 columns. The most straightforward way to analyse DRM data is to calculate means and variances of the repeatedly measured variables for each individual. For example, the average level of happiness of an individual felt over a day can be calculated by taking the average of the 5 happiness ratings. Similarly, the variance of these 5 ratings can be used to calculate a measure of how strongly happiness fluctuates. Using this type of individual-specific data allows us to explore an individual's

typical experience and can be used to show how people differ in these typical experiences. However, it does not show temporal or within-subject patterns. Figure 7 presents an example for this type of data structure.

Figure 7: Example of data structure 1

	id	sex	tsc	activityE1	happyE1	activityE2	happyE2
1	1	Female	41	Other	6	Commuting	6
2	2	Female	31	Watching TV	7	.	.
3	3	Female	27	Socializing	2	Other	3
4	4	Female	36	Exercising	6	.	.
5	5	Female	39	Taking care of kids	5	.	.
6	6	Female	26	Internet	4	Internet	5
7	7	Female	45	Watching TV	3	Watching TV	3
8	8	Female	29	Exercising	5	Watching TV	2
9	9	Female	35	Working/studying	5	Commuting	5
10	10	Female	41	Doing housework	6	Other	3
11	11	Male	48	Eating	7	Exercising	7
12	12	Female	44	Socializing	7	Socializing	7
13	13	Male	39	Resting/relaxing	5	Watching TV	4
14	14	Female	38	Doing housework	3	Commuting	2
15	15	Female	43	Resting/relaxing	2	Resting/relaxing	2

One row per episode

The data can also be structured on the episode level. In this case, each row represents one episode and the data from the same individual appears on multiple rows. For example, if an individual had recorded 5 episodes, the data from this individual appears in 5 rows and the happiness rating appears in one single column. This data structure allows us to investigate how context variables correlate with experiences, behaviours, and decisions. For example, happiness (within and across individuals) can differ according to:

- The location where an episode takes place.
- The activity conducted throughout the episode.
- The presence of other people during the episode.
- The visceral states such as hunger felt during the episode and so on.

For policy and business, this type of analysis is particularly valuable because some of the context-variables can be modified. For example, a new policy can change how tobacco products are displayed in supermarkets and that might affect individual behaviour. The DRM can help determine the antecedents, correlates, and also consequences of daily decisions and is an adequate tool to

better understand the context in which everyday decisions are made. Figure 8 presents example data where each row represents one episode.

Figure 8: Example of data structure 2

	id	sex	tsc	activityE	happyE
1	1	Female	41	Hobby*	6
2	1	Female	41	Commuting	6
3	1	Female	41	Eating	5
4	2	Female	31	Watching TV	7
5	3	Female	27	Socializing	2
6	3	Female	27	Other	3
7	3	Female	27	Resting/relaxing	3
8	4	Female	36	Exercising	6
9	5	Female	39	Taking care of kids	5
10	6	Female	26	Internet	4
11	6	Female	26	Internet	5
12	6	Female	26	Internet	7
13	7	Female	45	Watching TV	3
14	7	Female	45	Watching TV	3
15	7	Female	45	Watching TV	3

One row per unit of time

Naturalistic monitoring data can also be structured on the time level so that each row represents a time interval, such as an hour. In this case, each happiness rating can occur multiple times, depending on how long the respective episode lasted. For example, if an episode started at 3pm and ended at 5pm, and if the data is structured in 60-minute intervals, the same happiness rating will occur in two rows. Analyses based on the time interval can be used to measure the time courses of individual experiences. For example, the happiness of some individuals might increase, whilst that of others may decrease, over time. Preferences for food fluctuate throughout the day, levels of tiredness fluctuate from morning to night, demand for leisure differs between weekdays and the weekend. The effectiveness of interventions might increase or decrease over short time periods (and differently across sub-groups). The analyses of temporal dynamics lend itself to creating interesting graphs with time on the horizontal axis and other variables of interest on the vertical axis. Figure 9 presents example data where each row represents an hour for which one individual has provided information.

Figure 9: Example of data structure 3

	id	hourMid	sex	tsc	activityE	happyE
1	R_1IdRIUhrVuRBWDj	5	Male	39	Sleeping	4
2	R_5C0LHmvGOSu09Zb	6	Male	26	Watching TV	6
3	R_3BLhxwsW9fASwO9	6	Male	45	Working/studying	.
4	R_8Bt4g1UaWsfS7C5	7	Female	39	Sleeping	4
5	R_9zSnz2pYeWHQhcp	7	Female	54	Exercising	6
6	R_3OjgeXIcT1SJWW9	8	Female	33	Eating	7
7	R_6KlSoo02OuzusOB	8	Female	31	Taking care of kids	7
8	R_ePzP232dwywDtS5	8	Female	45	Sleeping	6
9	R_5isVKPaLIUEwb1H	8	Female	49	Personal care*	3
10	R_3lwTI6GfE1XhcEJ	8	Female	45	Personal care*	6
11	R_5hVOvZXvj4N24Ad	8	Female	56	Eating	2
12	R_9B9UJTTsQtcGYAt	8	Male	46	Taking care of kids	7
13	R_9pDnI1OKpZt9jh3	8	Male	27	Taking care of kids	5
14	R_5zkLCzSB57sa353	8	Male	39	Preparing food	1
15	R_7P2ygudV1QRut6d	8	Female	51	Eating	7

Multilevel-modelling

Whenever the same individual appears in more than one row, multilevel models to analyse naturalistic monitoring data are appropriate. These models take into account that observations of the same individual are not independent on each other. Most naturalistic monitoring studies ask the same participant to respond to the same question multiple times. It is likely that the answers from one participant correlate with each other more than the answers given by different participants. In order to take account for this “nested” data-structure, the analyses have to use multi-level models.

Multi-level models can also take into account that observations are ordered in time so that those close in time are more similar than those more distant in time. Due to this ordered nature of the data, it is important to control for time even if the research question does not refer to time trends, and to model autocorrelation in within-person error terms acknowledging that subsequent observations are similar to each other, which produces a downward bias in standard errors. Moreover, multilevel models can deal with unbalanced datasets in which the number of observations per individual differ across individuals. Another consideration the researcher has to take into account is whether independent variables should be separated into level 1 and level 2 components by person-mean cantering or group-mean cantering (i.e., calculating each individuals’ mean of the independent variable and subtracting it from the individuals raw score on the IV). Several textbooks are available that researchers should study when designing DRM studies and analysing DRM data.[19]

5 EXAMPLE RESULTS FROM A PREVIOUS STUDY

This section presents some descriptive statistics from a study that used an abbreviated online version of the DRM that we distributed in a nationally representative sample in Ireland. In the abbreviated DRM, participants were told that we are interested in what happened "yesterday" at some point in time that was randomly chosen between 8am and 9pm. The study then guided the participants through two steps as is typical in DRM studies. Participants could recall up to three episodes and in the second step, we went through each of the episodes chronologically, showed them the notes they had written down before, and asked them several follow-up questions. The sample (n=955) was representative of the Irish population based on gender, age, county of residency, and occupation. All participants together reported on 2667 episodes. Below, we present a selected subset of the associations we found in the data in order to illustrate some ways of analysing and presenting results from DRM studies. We focus on differences in everyday tiredness and how age and a measure of individual-specific trait self-control are associated with these differences.

5.1 Analysis on the individual level

First, we analyse the data at the individual level. In order to produce Figure 10 and Figure 11, we calculated the average tiredness over all episodes for each individual (on the vertical axis) and correlated this average with age and a measure of how self-controlled people perceive themselves to be in both figures, respectively. The figures show that older and more self-controlled people tended to be less tired in their daily lives.

Figure 10: Correlation between average tiredness and age.

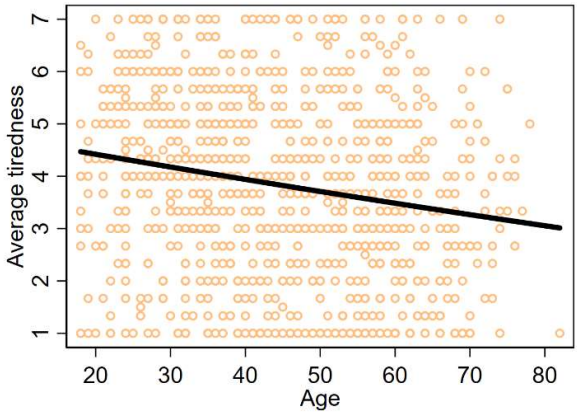
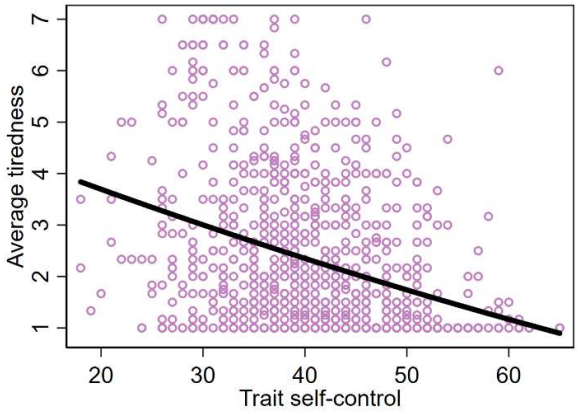


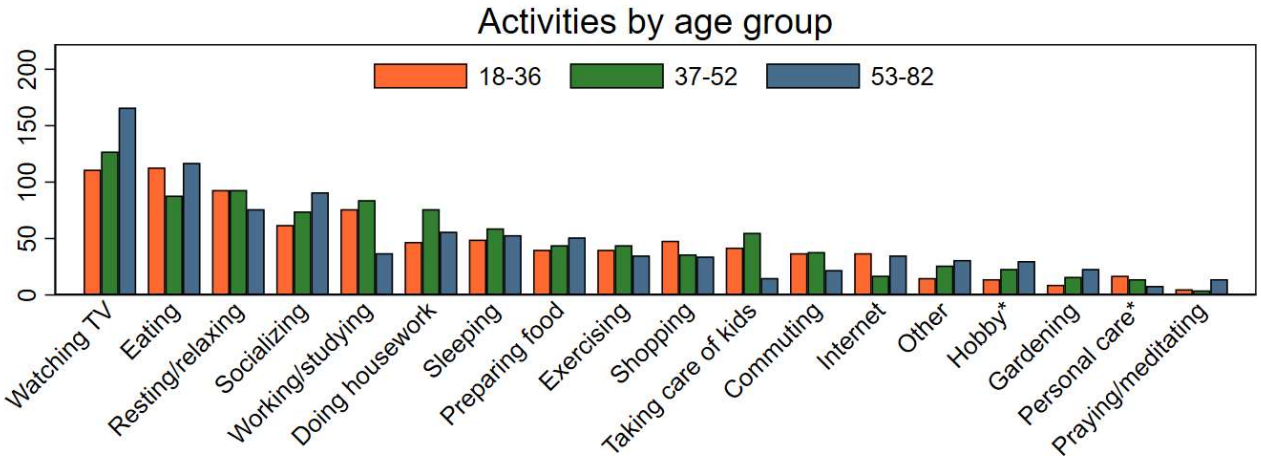
Figure 11: Correlation between average tiredness and trait self-control.



5.2 Analysis on the episode level

We can also analyse the data at the episode-level and Figure 12 present an example of this type of analyses. In order to produce the figure, we cut the age variable into three age segments and then counted how often people in each age segment engaged in each activity. The figure shows, for example, that watching TV was the most frequent activity, and that older participants watched more TV than younger participants.

Figure 12: Frequency of different activities enacted by different age groups.



5.3 Analysis on the time-unit analysis

For the final illustration, we reshaped the data to have a 10-minute interval in each row in the dataset. This allows a fine-grained temporal analysis as illustrated in the Figure 13 and Figure 14. Figure 13 shows that for people over 53, tiredness follows a U-shape over the day. These participants were as tired as younger participants in the mornings and the evenings, but throughout the day they were less tired. Figure 14 shows tiredness over the day for three different groups of participants, with low, medium, and high trait self-control. We used -1 standard deviation and +1 standard deviation as the cut off points for these groups. This figure shows that self-control matters for tiredness, but mainly at the start of the day and not later on.

Figure 13: Tiredness over the day for different age groups.

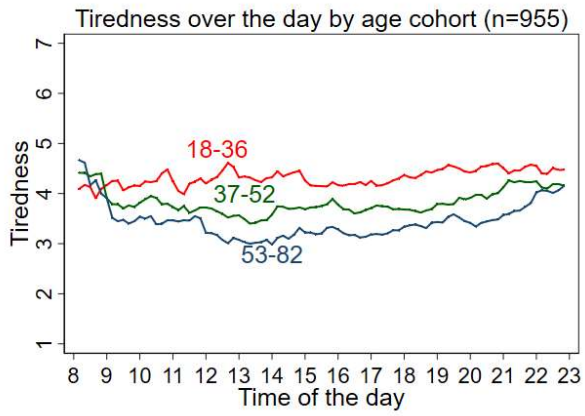
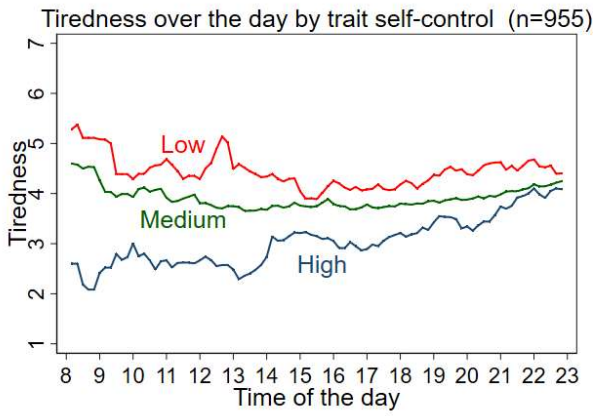


Figure 14: Tiredness over the day by trait self-control.



6 BUSINESS AND POLICY APPLICATIONS

6.1 The DRM and measuring everyday happiness

The DRM has been used to evaluate policy interventions, and recent research shows that it is possible to causally identify the effects of policy interventions on everyday measures of subjective

well-being.[22] The idea that subjective well-being data can and should be used to evaluate policies is receiving increasing attention.[23] The argument is that a policy can be evaluated according to whether and how much it changes happiness levels as people experience them in their daily lives.[24] If a policy increases positive affect and reduces negative affect, it is better than an alternative policy that does not do so. The DRM can help providing everyday subjective well-being data and thus contributes to the argument that an exclusive focus on financial indicators to measure human welfare is insufficient.[4]

6.2 The DRM and behavioural policy-making

Recent years have seen a rise in the application of findings from the behavioural sciences in business and policy contexts. All over the world, behavioural insights teams are being set up and governments as well as the private sector have begun to change behaviour using the most recent behavioural science findings.[25–28] In Ireland, behavioural insights are increasingly being applied systematically in both the public and private sector. Behavioural science, as it has come be understood recently, is the combination of behavioural economics and psychology. It builds on the evidence that people are not always rational decision-makers. Due to (predictable) deviations from rationality, people's decisions can often influenced by supposedly irrelevant factors of the "choice architecture", i.e. the context in which decisions are made.[28] Behavioural public policy deals with the intentional modification of the choice architecture in order to change behaviour. In its most common form, "nudging", behavioural public policies can often be implemented rather easily (e.g., compared to education campaigns and to regulations). Nevertheless, nudges can change behaviour significantly. Maybe because nudges do not reduce freedom of choice, they have gained considerable interest among politicians for reasons related to cost-effectiveness and ethics.[i]

Naturalistic monitoring research can play an important role for behavioural public policy-making. Many decisions that behavioural policies aim to change are everyday decisions and are affected by the everyday choice architecture. Naturalistic monitoring tools such as the DRM can provide detailed information about everyday choice architecture and contemporaneous decisions.[6] In particular, DRM research can (i) be used to improve the ecological validity of behavioural science studies, (ii) provide mechanistic evidence of the everyday workings of behavioural interventions, and (ii) help us to better understand people's true preferences. Many behavioural policy interventions focus on changing the everyday choice architecture. However, behavioural science research has produced more evidence on low-frequency decision-making than on everyday decisions. Using naturalistic monitoring research to evaluate everyday choice architecture will be

[i] For a guide to ethical nudging, see the FORGOOD framework REF.

a welcome addition to the behavioural public policy-makers toolbox in order to help identifying the mechanisms and temporal dynamics of everyday behavioural interventions.

7 CONCLUSION AND COLLABORATION

This brief suggested that the DRM can be a valuable tool to evaluate policies that aim to change behaviours in everyday life. The key benefits of the DRM, which set it apart from alternative approaches, are that it allows us to measure experiences, behaviours, and choice architecture in naturalistic, everyday contexts, and furthermore that it can be used in large, representative samples. It is thus a method that can help to quantify the extent to which policies change our behaviour in the real world.

The aim of the brief was to facilitate the use of the DRM in business and policy contexts in order to complement surveys, observational data, lab experiments, and randomized controlled trials. Therefore, the brief presented a guide to conducting Day Reconstruction Method (DRM) studies.

There is considerable potential in the DRM for a large number of applications in research and applied settings. Conducting research where it actually matters can provide many synergies and additional value for researches, policy-makers, and businesses by developing relevant study designs from a coherent, consistent, and credible analytical base.

If you are interested in collaborating on a DRM study, please contact the authors at admin@EnvEcon.eu.

REFERENCES

1. Deaton, A. & Cartwright, N. Understanding and misunderstanding randomized controlled trials. *Social Science & Medicine* **210**, 2–21 (2018).
2. Grüne-Yanoff, T. Why behavioural policy needs mechanistic evidence. *Economics and Philosophy* 1–21 (2015).
3. Shiffman, S., Stone, A. A. & Hufford, M. R. Ecological momentary assessment. *Annual Review of Clinical Psychology* **4**, 1–32 (2008).
4. Kahneman, D., Krueger, A. B., Schkade, D. A., Schwarz, N. & Stone, A. A. A survey method for characterizing daily life experience: The day reconstruction method. *Science* **306**, 1776–1780 (2004).
5. Delaney, L. & Lades, L. K. Present Bias and Everyday Self-Control Failures: A Day Reconstruction Study. *J. Behav. Dec. Making* (2017). doi:10.1002/bdm.2031
6. Lades, L. K., Martin, L. & Delaney, L. Informing behavioural policies with data from everyday life. *Behavioural Public Policy* 1–19 (2019). doi:10.1017/bpp.2018.37
7. Daly, M., Baumeister, R. F., Delaney, L. & MacLachlan, M. Self-control and its relation to emotions and psychobiology: evidence from a Day Reconstruction Method study. *Journal of Behavioral Medicine* **37**, 81–93 (2014).
8. Anusic, I., Lucas, R. E. & Donnellan, M. B. The Validity of the Day Reconstruction Method in the German Socio-economic Panel Study. *Soc Indic Res* **130**, 213–232 (2017).
9. Krueger, A. B., Kahneman, D., Schkade, D., Schwarz, N. & Stone, A. A. National time accounting: The currency of life. in *Measuring the subjective well-being of nations: National accounts of time use and well-being* 9–86 (University of Chicago Press, 2009).
10. Bakker, A. B., Demerouti, E., Oerlemans, W. & Sonnentag, S. Workaholism and daily recovery: A day reconstruction study of leisure activities. *Journal of Organizational Behavior* **34**, 87–107 (2013).
11. Daly, M., Delaney, L., Doran, P. P., Harmon, C. & MacLachlan, M. Naturalistic monitoring of the affect-heart rate relationship: A day reconstruction study. *Health Psychology* **29**, 186–195 (2010).
12. Ishio, J. & Abe, N. Measuring Affective Well-Being by the Combination of the Day Reconstruction Method and a Wearable Device: Case Study of an Aging and Depopulating Community in Japan. *Augment Hum Res* **2**, 2 (2017).
13. Knabe, A., Rätzel, S., Schöb, R. & Weimann, J. Dissatisfied with Life but Having a Good Day: Time-use and Well-being of the Unemployed. *The Economic Journal* **120**, 867–889 (2010).
14. Lee, P. H., Tse, A. C. Y. & Lee, K. Y. A new statistical model for the Day Reconstruction Method. *International Journal of Methods in Psychiatric Research* **26**, e1547 (2017).
15. Diener, E. & Tay, L. Review of the day reconstruction method (DRM). *Social Indicators Research* **116**, 255–267 (2014).
16. Dockray, S. *et al.* A Comparison of Affect Ratings Obtained with Ecological Momentary Assessment and the Day Reconstruction Method. *Social Indicators Research* **99**, 269–283 (2010).
17. Kim, J., Kikuchi, H. & Yamamoto, Y. Systematic comparison between ecological momentary assessment and day reconstruction method for fatigue and mood states in healthy adults. *British Journal of Health Psychology* **18**, 155–167 (2013).
18. Sonnenberg, B., Riediger, M., Wrzus, C. & Wagner, G. G. Measuring time use in surveys – Concordance of survey and experience sampling measures. *Social Science Research* **41**, 1037–1052 (2012).
19. Bolger, N. & Laurenceau, J.-P. Intensive Longitudinal Methods: An Introduction to Diary and Experience Sampling Research. (Guilford Press, 2013).
20. University of Bristol. Centre for Multilevel Modelling | Centre for Multilevel Modelling | University of Bristol. Available at: http://www.bristol.ac.uk/cmm/. (Accessed: 17th May 2019)
21. Grube, A., Schroer, J., Hentzschel, C. & Hertel, G. The event reconstruction method: An efficient measure of experience-based job satisfaction. *Journal of Occupational and Organizational Psychology* **81**, 669–689 (2008).
22. Doyle, O., Delaney, L., O'Farrelly, C., Fitzpatrick, N. & Daly, M. Can Early Intervention Improve Maternal Well-Being? Evidence from a Randomized Controlled Trial. *PLOS ONE* **12**, e0169829 (2017).
23. Odermatt, R. & Stutzer, A. Subjective Well-Being and Public Policy. in *Handbook of Well-Being* (eds. Diener, E., Oishi, S. & Tay, L.) (DEF Publishers, 2017).
24. Layard, R. Happiness and Public Policy: a Challenge to the Profession. *The Economic Journal* **116**, C24–C33 (2006).
25. OECD. Behavioural Insights and Public Policy - Lessons from Around the World - en - OECD. (2017). Available at: http://www.oecd.org/gov/regulatory-policy/behavioural-insights-and-public-policy-9789264270480-en.htm. (Accessed: 18th October 2018)
26. Jolls, C., Sunstein, C. R. & Thaler, R. H. A behavioral approach to law and economics. *Stanford law review* 1471–1550 (1998).
27. Halpern, D. The Rise of Psychology in Policy: The UK's de facto Council of Psychological Science Advisers. *Perspectives on Psychological Science* **10**, 768–771 (2015).
28. Thaler, R. H. & Sunstein, C. R. *Nudge: Improving decisions about health, wealth, and happiness.* (Yale University Press, 2008).

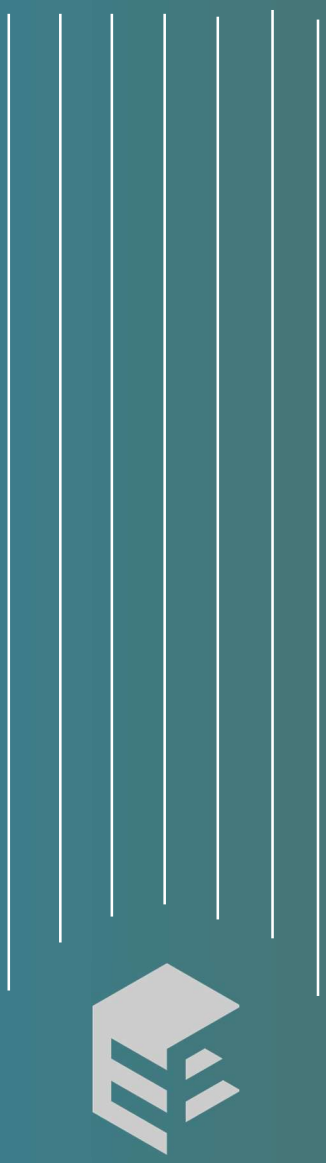
Contacts for more information

Dr. Leonhard K. Lades

Assistant Professor, UCD
Environmental Policy

Dr. Andrew Kelly

CEO, EnvEcon



Stay in touch

https://envecon.eu/

admin@EnvEcon.eu

EnvEcon

11 Priory Office Park,
Stillorgan Road,
Blackrock, Co.
Dublin, Ireland