



Contents lists available at ScienceDirect

## Journal of Mathematical Psychology

journal homepage: [www.elsevier.com/locate/jmp](http://www.elsevier.com/locate/jmp)

Theoretical Note

## Two peas in a pod: Discounting models as a special case of the VARMAX

Niels Vanhasbroeck\*, Tim Loossens, Francis Tuerlinckx

Research Group of Quantitative Psychology and Individual Differences, KU Leuven, Leuven, Belgium

## ARTICLE INFO

## Keywords:

Exponential discounting  
Autoregression  
Computational model  
Dynamical model  
Quasi-hyperbolic discounting

## ABSTRACT

In this paper, we establish a formal connection between two dynamic modeling approaches that are often taken to study affect dynamics. More specifically, we show that the exponential discounting model can be rewritten to a specific case of the VARMAX, thereby shedding light on the underlying similarities and assumptions of the two models. This derivation has some important consequences for research. First, it allows researchers who use discounting models in their studies to use the tools established within the broader time series literature to evaluate the applicability of their models. Second, it lays bare some of the implicit restrictions discounting models put on their parameters and, therefore, provides a foundation for empirical testing and validation of these models. One of these restrictions concerns the exponential shape of the discounting function that is often assumed in the affect dynamical literature. As an alternative, we briefly introduce the quasi-hyperbolic discounting function.

## 1. Introduction

Psychological processes unfold over time, implying that we need to actively consider the temporal nature of these processes to get a complete understanding of them. Statistical or computational models are a popular way to study these time-dependent fluctuations, given that they can serve as a formal basis for understanding and investigating the key dynamical features of a process of interest. One area of research that has embraced these computational models is the field of *affect dynamics* – which revolves around how emotional feelings change over time (Waugh & Kuppens, 2021) – as evidenced by the many computational models that are used within this field, going from linear to nonlinear, from discrete- to continuous-time, and from descriptive to theoretical models (to give a few examples, Ariens et al., 2020; Bennett et al., 2022; Deboeck & Bergeman, 2013; Loossens et al., 2020; Rutledge et al., 2014; Steele & Ferrer, 2011; Voelkle & Oud, 2013). Unfortunately, however, it is often unclear how these models relate to each other, which may unintentionally contribute to difficulties with the adoption of these models in substantive research. This currently limits the extent to which the promise of these models can be fulfilled.

Many dynamical models are related to each other, although the nature of these relationships is not always obvious. It is therefore of great importance to document the relationships between the different computational models that are used in the literature. In this article, we will investigate the relationship between two classes of models that are frequently used in the affect dynamical literature, namely the

*discounting models* and the *autoregressive models*, both of which we will introduce in the next sections.

Let us first establish the mathematical notation we use in this text. Variables that are observed or measured are denoted with Roman letters, while Greek letters are reserved for the model parameters. Lowercase letters are used for scalars or, when they are bold, for vectors. Uppercase letters are used for matrices.

## 1.1. Discounting models

Despite the long presence of the discounting model within the economic literature (e.g., Cairns & van der Pol, 2000; Koopmans, 1960; Myerson & Green, 1995), Rutledge et al. (2014) only recently introduced it to the affect dynamical literature. In their study, Rutledge et al. (2014) investigated the relationship between participant's happiness and the monetary outcomes that they received in a gambling paradigm. This paradigm required participants to repeatedly choose either a certain outcome or an uncertain gamble between two equally probable outcomes. If the participant selected the certain outcome, the reward was immediately added to the participant's total. If the gamble was chosen, one of the outcomes was selected at random and was then added to the participant's total.

To formalize the relationship between happiness and the outcomes of their experiment, Rutledge et al. (2014) created the following computational model:

$$h_t = \alpha + \beta_1 \sum_{j=0}^{t-1} \gamma^j cr_{t-j} + \beta_2 \sum_{j=0}^{t-1} \gamma^j ev_{t-j} + \beta_3 \sum_{j=0}^{t-1} \gamma^j rpe_{t-j} + \epsilon_t, \quad (1)$$

\* Correspondence to: Tiensestraat 102, 3000 Leuven, Belgium.  
E-mail address: [niels.vanhasbroeck@gmail.com](mailto:niels.vanhasbroeck@gmail.com) (N. Vanhasbroeck).

where

$$\epsilon_t \sim N(0, \sigma^2),$$

and where  $h_t$  was a person's self-reported happiness at trial  $t$ ,  $cr$  is the value of the certain outcome (if chosen),  $ev$  is the expected value of the gamble outcomes (if chosen), and  $rpe$  is the difference between the  $ev$  and the actual obtained gamble outcome. In this equation, the happiness  $h$  at time  $t$  is a function of the cumulative sum of the discounted previous outcomes  $cr$ ,  $ev$ , and  $rpe$ , meaning that the history of the received outcomes matters in determining one's current happiness. To ensure that recent outcomes have a greater influence on happiness, Eq. (1) discounts past outcomes exponentially over time, the strength of which is determined by the parameter  $\gamma \in [0, 1)$  which we will refer to as the *discounting factor*.

Eq. (1) closely resembles one of the solutions to the equation of the *expected return*, a concept that was introduced in the reinforcement learning literature. Expected return is the total reward people expect to receive at the end of a task if they behave in a given way and can be formalized as a function of all future rewards (Sutton & Barto, 2018):

$$e_t = f(r_{t+1}, \dots, r_{t+n}), \quad (2)$$

where  $e$  is the expected return looking forward from the current moment  $t$ ,  $r$  is the reward that the agent receives at a given future time  $t+j$ , and  $f$  is a function of these future rewards. For the model to describe a stable process, the function  $f$  should converge to a finite value, even when the task itself continues forever (i.e., when  $n \rightarrow \infty$ ). Because of this,  $f$  is typically assumed to take on a discounting form, such as (see Sutton & Barto, 2018):

$$e_t = \sum_{j=0}^{\infty} \gamma^j r_{t+j+1}. \quad (3)$$

Note that this discounted sum is the same as the sum in Eq. (1) with one important difference, namely the frame of reference. While Eq. (3) deals with the discounting of future rewards, Eq. (1) is concerned with the discounting of outcomes that one has already received. One model thus deals with the future while the other one deals with the past. Despite this difference in the frame of reference, the discounting model itself remains the same.

At this point, it is also useful to make the connection to research around *delay discounting*, which is concerned with how people discount rewards that are received at different moments in the future (see e.g., Ballard et al., 2023; Blavatsky, 2016; Myerson & Green, 1995).<sup>1</sup> Within this field, Samuelson (1937) proposed the exponential discounting function, which has a similar functional form to those in Eqs. (1) and (3) and which has since received ample attention. However, the exponential discounting model is not uncontroversial and some alternatives have been proposed. In the first section of this paper, we will only concern ourselves with relating the discounting model as used in the literature to the autoregressive models. Later, we will point out the need to empirically validate the exponential shape of the discounting function and propose the quasi-hyperbolic model as an alternative (Laibson, 1997).

In this article, we will use a generalized version of Eq. (1) which we define as (see Appendix A)<sup>2</sup>:

$$y_t = \alpha + \sum_{j=0}^{t-1} \Gamma^j B x_{t-j} + \epsilon_t \quad (4)$$

$$\epsilon_t \sim N(\mathbf{0}, \Sigma),$$

where the  $d$ -dimensional vector  $y_t$  contains the values for the dependent variables at time  $t$ . The  $d$ -dimensional vector  $\alpha$  represents the

intercept, which coincides with the mean of the process  $y$ . Within this sum, the  $d \times d$  matrix  $\Gamma$  contains the discounting factors, defining the extent to which previous values of  $x$  influence the current values of  $y$ . Importantly,  $\Gamma$  is subject to several restrictions. First, the moduli of the eigenvalues of  $\Gamma$  are restricted to be smaller than 1, which ensures that the sum converges to a stable value (cf. restriction on function  $f$ , Sutton & Barto, 2018).<sup>3</sup> Second, the values in  $\Gamma$  are restricted to be positive (Rutledge et al., 2014; Sutton & Barto, 2018), which means that the discounting of a value over time does not overshoot the baseline state (cf. the sawtooth pattern in Fig. 2). Finally, the  $k$  independent variables are contained in the vector  $x_t$  and get scaled by the  $d \times k$  matrix  $B$ . The matrix  $B$  determines the strength of the association between the discounted sum of the independent variables with each of the variables in  $y$ , playing the same role as the  $\beta$ 's in Eq. (1). We refer the reader to Fig. 1 for a visual representation of what each parameter does in the model.

After its introduction in the field of affect dynamics, the discounting model has been widely used to understand how affect changes in relation to its environment, both by the original authors (e.g., Blain & Rutledge, 2020; Chew et al., 2021; de Berker et al., 2016; Keren et al., 2021; Rutledge et al., 2017, 2015) and by independent researchers (e.g., Asutay & Västfjäll, 2022; Vanhasbroeck et al., 2021; Villano et al., 2020; Vinckier et al., 2018). Given its rising influence, it is important to find out how this model relates to other, already existing models within the field.

### 1.2. Autoregressive models

The autoregressive model is another model popular in the emotion literature, usually serving as a data-analytic tool to investigate how affect changes over time (e.g., Adolf et al., 2017; Booij et al., 2018; Congard et al., 2011; Hamaker et al., 2005; Sperry et al., 2020). While several variants of the autoregressive model exist, what binds these models together is the shared assumption that current values of a dependent variable  $y$  are related to past values of the same variable, implying that values of  $y$  are carried over across time. Formally, this assumption is implemented by regressing past values of a variable to its current value, so that:

$$y_t = \delta + \Theta y_{t-1} + \epsilon_t, \quad (5)$$

where

$$\epsilon_t \sim N(\mathbf{0}, \Sigma).$$

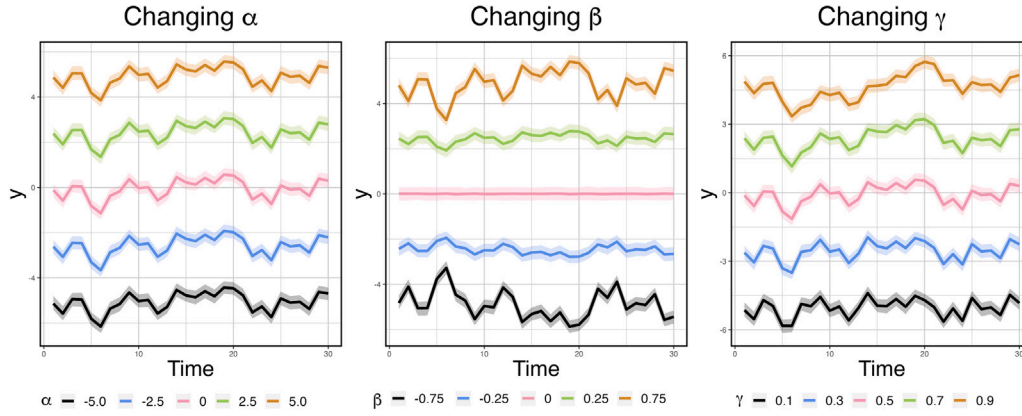
This model is known as the *vector autoregressive model* or VAR. The  $d$ -dimensional vector  $\delta$  contains the intercepts. The next term is the autoregressive component, which scales previous values of  $y$  by the  $d \times d$  transition matrix  $\Theta$ . This component defines an exponential decay of values of  $y$  towards the baseline state  $\mu$ , which is equal to  $(I_d - \Theta)^{-1} \delta$  with  $I_d$  being the  $d \times d$  identity matrix. Importantly, the moduli of the eigenvalues of  $\Theta$  should be smaller than 1 in order for the model to describe a stable process (Hamilton, 1994, also cf.  $\Gamma$ ).

As we will show in the next section, the discounting model is a special case of a slightly more complicated autoregressive model, namely the *moderated vector autoregressive moving-average model* or VARMAX (Adolf et al., 2017; Shumway & Stoffer, 2006). Like the name suggests, this model consists of an autoregressive component, a moving

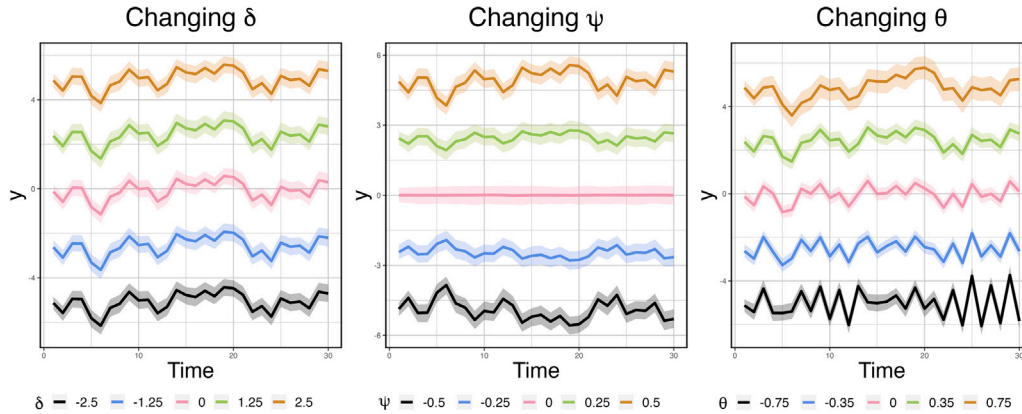
<sup>3</sup> Note that when the multidimensional discounting model of Eq. (4) was used by researchers, they restricted the discounting matrix  $\Gamma$  to be diagonal (e.g. Asutay & Västfjäll, 2022; Villano et al., 2020). This restriction implies that the discounting of  $x$  is independent for each of the variables in  $y$ , or that there is no dynamical relation between the different dependent variables. Given that no such restriction is necessary for our discussion, we allow  $\Gamma$  to have off-diagonal elements.

<sup>1</sup> We thank the reviewers for making this connection.

<sup>2</sup> Note that in both Eqs. (4) and (6) the  $d \times d$  matrix  $\Sigma$  defines the (co)variances of the residuals. Given that  $\Sigma$  does not directly influence our results, we do not discuss it in text.



**Fig. 1.** Visualization of the role each parameters plays in the one-dimensional discounting model. These plots display the median and 95% confidence intervals of 1000 simulated time series of length 30, where  $y$  depended on the same set of values for  $x$  (drawn from a standard normal distribution). We systematically varied the values of the model that was used to simulate the time series in this figure, allowing the parameters to take on the following values:  $\alpha \in \{-5.00, -2.50, 0.00, 2.50, 5.00\}$ ,  $\beta \in \{-0.75, -0.25, 0.00, 0.25, 0.75\}$ ,  $\gamma \in \{0.10, 0.30, 0.50, 0.70, 0.90\}$ , and  $\sigma = 0.15$ . In each plot, we changed the values of  $\alpha$  to make the lines non-overlapping. The parameters' influence on the dynamics of the model can be summarized as follows. By changing the values of  $\alpha$ , the function is moved upwards or downwards, going together with higher or lower values of  $y$  respectively. Changing the values of  $\beta$  leads to changes in the direction and strength of the association between  $y$  and  $x$ . Finally, the values of  $\gamma$  control to which extent previous values of  $x$  carry over to the next time points, so that higher values of  $\gamma$  lead to a greater inertia in the time series.



**Fig. 2.** Visualization of the role the parameters play in the one-dimensional VARMAX. As shown in the next section, the parameters roughly correspond to those of the discounting model, and we will therefore not discuss the observed patterns in detail. However, we do bring the sawtooth pattern that is created by negative values of  $\theta$  to attention, as this is a pattern that cannot be created by the discounting model.

average component, and a set of external variables that are thought to influence  $y$ . Combined, this model then becomes:

$$y_t = \delta + \Psi x_t + \Theta y_{t-1} + \Phi \epsilon_{t-1} + \epsilon_t, \tag{6}$$

where again

$$\epsilon_t \sim N(\mathbf{0}, \Sigma).$$

In this equation,  $y$  changes with the values of the  $k$  independent variables  $x$ , which are scaled by the  $d \times k$  matrix  $\Psi$ . Additionally, fluctuations in  $y$  are captured with a moving-average component, which consists of the  $d \times d$  matrix  $\Phi$  and defines the extent to which residuals at the previous time point  $t - 1$  are carried over to the current observations of  $y$ . Importantly, the restriction on  $\Theta$  also applies to  $\Phi$ , which in this case makes sure that the model is *invertible* (Shumway & Stoffer, 2006).

Fig. 2 visualizes the effect of the parameters on  $y$ . When looking at this figure, it is useful to keep in mind that while most parameters map directly onto the discounting model, we should mention two differences with the time series shown in Fig. 1. First, the values of the discounting factors in  $\Gamma$  are restricted to be positive, a restriction that does not apply to the transition matrix  $\Theta$ . This allows the autoregressive models to capture a sawtooth-like dynamical pattern, as shown in Fig. 2. Second, changing the value of  $\Theta$  independently of  $\delta$  also changes the value of the baseline  $\mu$ . To keep all time series equidistant, we

counteracted this effect by changing the values of  $\delta$  simultaneously with the values of  $\theta$ . The full effect of only changing  $\theta$  is thus not shown in this figure.

Due to their attractive properties and their intuitiveness, autoregressive models have been exceedingly popular in the affect dynamical literature. Typically, they are used as a data analytic tool (e.g., Dejonckheere et al., 2021; Lucas & Donnellan, 2007; Wendt et al., 2020), but autoregressive models have also been used to gain a deeper theoretical understanding of the dynamic structure of affect (e.g., Bonsall et al., 2012; de Haan-Rietdijk et al., 2017; Loossens et al., 2021).

## 2. Relating the models

We start our demonstration with the discounting model, of which the definition is repeated here:

$$y_t = \alpha + \sum_{j=0}^{t-1} \Gamma^j B x_{t-j} + \epsilon_t.$$

First, we take the most recent reward out of the sum (i.e., the reward for which  $j = 0$ ), so that this equation becomes:

$$y_t = \alpha + \Gamma^0 B x_{t-0} + \sum_{j=1}^{t-1} \Gamma^j B x_{t-j} + \epsilon_t$$

$$\begin{aligned}
 &= \alpha + I_d B x_t + \sum_{j=1}^{t-1} \Gamma^j B x_{t-j} + \epsilon_t \quad (\Gamma^0 = I_d) \\
 &= \alpha + B x_t + \sum_{j=1}^{t-1} \Gamma^j B x_{t-j} + \epsilon_t \quad (I_d B = B).
 \end{aligned}$$

Now, we factor  $\Gamma$  out of the sum and add the terms  $\Gamma\alpha - \Gamma\alpha$  and  $\Gamma\epsilon_{t-1} - \Gamma\epsilon_{t-1}$  in two subsequent steps, so that:

$$\begin{aligned}
 y_t &= \alpha + B x_t + \Gamma \sum_{j=1}^{t-1} \Gamma^{j-1} B x_{t-j} + \epsilon_t \quad (\text{Factor out } \Gamma) \\
 &= \alpha + \Gamma\alpha - \Gamma\alpha + B x_t + \Gamma \sum_{j=1}^{t-1} \Gamma^{j-1} B x_{t-j} + \Gamma\epsilon_{t-1} - \Gamma\epsilon_{t-1} + \epsilon_t.
 \end{aligned}$$

Next, we take together common terms by factoring out  $\Gamma$  and adjust the indices of the sum so that the sum starts at  $j = 0$ :

$$y_t = \alpha - \Gamma\alpha + B x_t + \Gamma \left( \alpha + \sum_{j=0}^{(t-1)-1} \Gamma^j B x_{(t-1)-j} + \epsilon_{t-1} \right) - \Gamma\epsilon_{t-1} + \epsilon_t.$$

Finally, we realize that the terms between the brackets are equal to  $y_{t-1}$ , and we can thus rewrite this equation to:

$$y_{t-1} = (I_d - \Gamma)\alpha + B x_t + \Gamma y_{t-1} - \Gamma\epsilon_{t-1} + \epsilon_t,$$

which is a special case of the VARMAX with its parameters equal to:

$$\begin{aligned}
 \delta &= (I_d - \Gamma)\alpha \\
 \Psi &= B \\
 \Theta &= \Gamma \\
 \Phi &= -\Gamma.
 \end{aligned}$$

This concludes our demonstration of the relationship between the two models. It is important to note that the result of our derivation adheres to the restrictions imposed on  $\Theta$  and  $\Phi$ , as  $\Gamma$  falls under the same restrictions. Furthermore notice that  $\alpha$  of the discounting model is equal to the baseline  $\mu$  of the VARMAX.

### 3. Properties

The main advantage of our demonstration is that it makes common assumptions of the discounting model and the VARMAX explicit while simultaneously exposing some of the discounting model's restrictions. On the one hand, this means that we can build upon the vast literature around time series models to evaluate the applicability of the discounting model on observed data. For example, one can use established tests of the stationarity assumption, which dictates that the statistical properties of the time series do not change over time (Scargle, 1981), and turn to solutions such as trend analysis or differencing in case the data are not stationary (see e.g. Hamilton, 1994; Shumway & Stoffer, 2006; Velicer & Molenaar, 2012). A wide range of tools becomes available to researchers who use the discounting model to fit their data.

On the other hand, we can leverage on the differences between the discounting model and the VARMAX to investigate the dynamical features of psychological processes. Specifically, one can reveal to which extent the assumptions apply to the process of interest by comparing the performance of these models on empirical data. For example, if two models share all assumptions except for one, and model comparison shows that one performs better than the other, this provides immediate evidence for or against the assumption that is not shared among the models. It is therefore important to be aware of the assumptions of a model and the behavior that these assumptions imply, as model comparison may reveal which of these apply to the process of interest.

Researchers should therefore empirically test the viability of the restrictions that come with the discounting model. Concretely, one can use data to estimate restricted and unrestricted versions of the VARMAX and use model comparison techniques to find out which type of model performs best. Testing these properties will provide us with

important information on the dynamical features of the system that we are interested in, enhancing our understanding of these systems.<sup>4</sup>

In what follows, we will discuss two restrictions that the discounting model imposes on the VARMAX. We furthermore discuss one property that is shared among the two models. Importantly, all three properties can be tested in the way that we described above and we encourage researchers to do so.

#### 3.1. Restriction 1: Single lag structure

The first restriction concerns the discounting model's single-lag in the autoregressive and moving-average components, meaning that the dependent variables only depend on their immediately predecesing value and residual. To appreciate this restriction, consider the generalized VARMAX that contains  $p$  lags in the autoregressive component and  $q$  lags in the moving average component (Hamilton, 1994; Shumway & Stoffer, 2006):

$$y_t = \delta + \Psi x_t + \sum_{j=1}^p \Theta_j y_{t-j} + \sum_{j=1}^q \Phi_j \epsilon_{t-j} + \epsilon_t. \tag{7}$$

This equation allows  $y$  at time  $t$  to be related to each of its predecesing values up to time point  $t - p$  and each of its predecesing residuals up to time point  $t - q$ . Interestingly, the discounting model restricts both  $p$  and  $q$  to be equal to 1, so that only the previous observation  $y_{t-1}$  and the previous residual  $\epsilon_{t-1}$  are related to the current observation  $y_t$ .

The discounting model thus makes a concrete prediction with regard to the structure of the data. More concretely, it predicts that a dynamical system should only show a temporal dependence of a single lag in both the autoregressive and moving-average components. This implies that whenever a dynamical system displays dependence of  $y$  on observations and/or residuals of two or more lags in the past, the discounting model cannot be the data generating model and might not be appropriate for use on these data. To test this restriction, one can compare the multilag VARMAX to a VARMAX that only assumes a single lag (Eq. (6) vs. Eq. (7)).

#### 3.2. Restriction 2: White noise

A second restriction of the discounting model dictates that the value of  $\Phi$  is equal to  $-\Theta$ , or the reverse of the transition matrix. This implies that residuals are not carried over across time, meaning that apart from the external factors in  $x$ , changes in  $y$  are only governed by white noise. To see this, consider what makes up the current values of  $y$  at time  $t$  under both models. Starting with the VARMAX, we can rewrite this formula in terms of several discounted sums (see Appendix B):

$$y_t = \mu + \sum_{j=0}^{t-1} \Theta^j \Psi x_{t-j} + \sum_{j=0}^{t-1} \Theta^j \epsilon_{t-j} + \sum_{j=1}^{t-1} \Theta^{j-1} \Phi \epsilon_{t-j}.$$

This equation shows that the fluctuations in  $y$  are a function of the discounted sums of the external factors  $x$  and the residuals  $\epsilon$  at time points  $t - j$ . In other words, the residuals are carried over across time and have a continuing influence on the process  $y$ .

Now consider what happens when we impose that  $\Phi = -\Theta$ , then this equation simplifies to:

$$\begin{aligned}
 y_t &= \mu + \sum_{j=0}^{t-1} \Theta^j \Psi x_{t-j} + \sum_{j=0}^{t-1} \Theta^j \epsilon_{t-j} - \sum_{j=1}^{t-1} \Theta^j \epsilon_{t-j} \\
 &= \mu + \sum_{j=0}^{t-1} \Theta^j \Psi x_{t-j} + \Theta^0 \epsilon_t + \sum_{j=1}^{t-1} \Theta^j \epsilon_{t-j} - \sum_{j=1}^{t-1} \Theta^j \epsilon_{t-j} \\
 &= \mu + \sum_{j=0}^{t-1} \Theta^j \Psi x_{t-j} + \epsilon_t,
 \end{aligned}$$

<sup>4</sup> It goes without saying that the strength of such a test depends on the characteristics of the study from which the test data were gathered.

which brings us back to the discounting model in Eq. (4). Unlike the typical VARMAX, the residuals are thus not carried over across time, meaning they are not a part of the process  $y$ . In other words, when  $\Phi = -\theta$  the VARMAX assumes that the residuals represent white noise. The discounting model thus expects temporal changes in  $y$  to only come from the external variables  $x$ .

While this restriction may not seem that important, it does have important interpretational consequences. Specifically, the discounting model assumes that residuals are independent of the process itself, and may thus be interpreted as process-independent error. For the typical autoregressive models, the residuals have a lasting impact on the system and are therefore interpreted as a (meaningful) part of the system. Typically, these residuals are assumed to reflect accumulated sum of internal and external, mutually independent events that partially determined to the observed state of the system (Hamaker, 2012; Kuppens et al., 2010; Loossens et al., 2020). Whatever the source of this error, the VARMAX assumes that it is inherent to  $y$  while the discounting model assumes the error is independent of the process.

To test this restriction, one has to first estimate a VARMAX that restricts the value of  $\Phi$  to  $-\theta$  and a VARMAX that allows for the free estimation of  $\Phi$ . Then, the performance of these two models should be compared to find out whether this restriction holds up against the data.

### 3.3. Property 1: Exponential decay

One important property of the discounting model as it is currently defined is the exponential shape of the discounting function. As we mentioned before, the exponential discounting function has been a topic of investigation in the delay discounting literature, where evidence against this function has accumulated over the years. Instead, researchers in this field are turning to other kinds of discounting functions, such as the hyperbolic, double exponential, and additive discounting functions (Ballard et al., 2023; Mazur, 1987; van den Bos & McClure, 2013). Similarly, we might question the validity of the exponential decay functions used in many linear dynamical models such as the discounting model and the VARMAX.

To investigate this question, one should first define other kinds of discounting models with which the exponential one can be compared. For this, researchers can use the approach taken by Rutledge et al. (2014) in their paper. Specifically, one can select a discounting function, plug it into a cumulative sum, and derive the recursive counterpart of this model. This way, competing models for the typical exponential discounting model are created, allowing one to test its relative fit to empirical data. It will furthermore bring the literature around discounting models closer to the literature around other models that rely on recursion, such as autoregressive models and reinforcement learning models.

As an example, we propose the quasi-hyperbolic discounting model as an alternative to exponential discounting. In one dimension, the quasi-hyperbolic discounting model is defined as (see Laibson, 1997):

$$y_t = \alpha + \beta \sum_{j=0}^{t-1} \kappa^{i(j)} v^j x_{t-j} + \epsilon_t,$$

where  $\kappa, v \in [0, 1)$  and where  $i(t)$  is an indicator function so that:

$$i(t) = \begin{cases} 0 & \text{if } t = 0 \\ 1 & \text{if } t \in \{1, 2, \dots\} \end{cases},$$

In multiple dimensions, this model can be defined as:<sup>5</sup>

$$y_t = \alpha + \sum_{j=0}^{t-1} N^j K^{i(j)} B x_{t-j} + \epsilon_t, \tag{8}$$

<sup>5</sup> Note that one can define an alternative quasi-hyperbolic model in which the multiplication of  $N$  and  $K$  switch places, so that the sum becomes  $\sum_{j=0}^{t-1} K^{i(j)} N^j B x_{t-j}$ . The derivation of this model differs in its result from the one explained in-text.

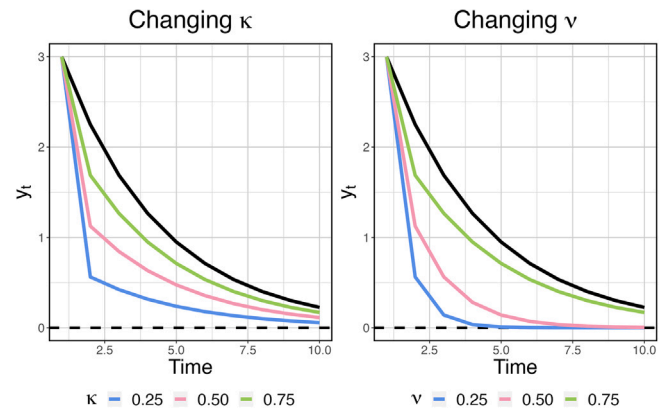


Fig. 3. Visualization of the expected regulation patterns of the quasi-hyperbolic discounting model versus the exponential discounting model (black solid line). In the left panel, the effect of varying  $\kappa$  is shown while keeping the discounting factors of the quasi-hyperbolic  $v$  and of the exponential model  $\gamma$  equal to 0.75. While starting out at the same values,  $\kappa$  determines how strongly the regulation is going from the first to the second time point. After this initial effect, the regulatory pattern follows an exponential curve defined by  $v$ . In the right panel, the effect of varying  $v$  is shown while fixing  $\kappa = \gamma = 0.75$ . As one might suspect,  $v$  quantifies the strength of the exponential regulation towards the baseline.

where the  $d \times d$  matrices  $K$  and  $N$  conform to the same restrictions as  $\Gamma$ . This model can be rewritten to an autoregressive model of the form (see Appendix C):

$$y_t = (I_d - N)\alpha + Bx_t + N(K - I_d)Bx_{t-1} + Ny_{t-1} - N\epsilon_{t-1} + \epsilon_t,$$

which corresponds to a VARMAX with an additional lag in the predictor variables  $x$ .

Fig. 3 visualizes how the expected regulation pattern for the unidimensional quasi-hyperbolic discounting model compared to the exponential one. As one may see, the parameter  $K$  controls the extent to which there is a “recency bias”, which entails an automatic devaluation of previous values of  $x$  because they are in the past. This creates a sharp drop in the regulation patterns shown in Fig. 3. Once this sharp drop has been introduced, the regulation patterns follows an exponential distribution as defined by the parameter  $N$ . Interestingly, this implies that when  $K = I_d$ , the quasi-hyperbolic discounting model is equal to the exponential discounting model. In other words, the exponential discounting model is nested within the quasi-hyperbolic discounting model.

## 4. Discussion

In this paper, we demonstrated the relationship between two classes of computational models that are frequently used to study affect dynamics. Specifically, we demonstrated that the (exponential) discounting model is nested within the autoregressive models, more specifically as a special case of the VARMAX. This relationship shows that the discounting model and the VARMAX describe a very similar process, specifically one in which the fluctuations in the variables  $y$  are explained by a discounted sum of all past values of the predictor variables  $x$ . For affect dynamics, this means that affective fluctuations are a manifestation of all external events that an individual encountered throughout the past.

This is an interesting observation, as researchers typically interpret the results that come from both models in a different way. Autoregressive models are generally interpreted as capturing the dynamics of the process of interest in relation to external factors. However, it is still believed that the observed dynamics are intrinsic to the studied process. In other words, external factors may perturb the system, but it is the system itself that recovers from such perturbations (for a similar perspective, see Kuppens & Verduyn, 2015). Discounting models on

the other hand are interpreted as showing that the process of interest solely arises through a specific set of external factors of which the influence decreases over time. In this view, the studied process is thus a “playing ball” of the external influences and, consequently, lacks its own intrinsic dynamical features.

The fact that discounting models are nested within the VARMAX shows that the divergence between the interpretations may be ill-advised. Discounting models imply a given dynamical structure that cannot be easily disentangled from the autoregressive structure of the VARMAX, even when the restrictions above do not hold. One should thus be careful when interpreting the results that come from applying a discounting model to psychological variables such as affect.

Our derivation clarifies the position of the discounting model within the literature on time series analysis, allowing users to draw on this literature to test and deal with violations of common assumptions such as stationarity. We also reconnected the discounting model with the reinforcement learning literature and, more importantly, the literature around delay discounting. However, one should consider an important difference between the models used in these bodies of literature and the model that was introduced by Rutledge et al. (2014): While the model in Eq. (4) is concerned with discounted sums of outcomes in the past, the discounting function that it builds upon was originally introduced to capture how people discount the values of future rewards (and similar for Eq. (8)). There is thus a different frame of reference for how the discounting functions were originally used and how we use them in this article.

Given the different frame of reference, one may question the applicability of future-oriented discounting functions to understand the decaying patterns in variables such as affect. This concern is valid, and in no way do we advocate for the blind use of discounting models to investigate affect dynamics. We instead hope that we gave researchers the tools to investigate the dynamical structure of their data by making the exponential discounting assumption explicit and by proposing an alternative model. Our approach is thus a purely pragmatic one.

Our derivation also exposes some of the restrictions the discounting model imposes on the VARMAX, for which we discussed several areas of future research, which we repeat here. First, research should systematically investigate whether affect evolves only in relation to its immediate predecesing value and residual, or in relation to more distant lags as well (Restriction 1). If a single lag suffices, than the models that are discussed in this article are appropriate for use on empirical data. If on the other hand more than one lag should be accounted for, then researchers should employ models that allow for more than one lag to understand our data (e.g., Eq. (7)). Testing this restriction will thus reveal a critical aspect of the dynamical structure of affect.

Second, future research should investigate whether the parametric restriction that the discounting model imposes on the moving average component of VARMAX holds up against empirical data ( $\Phi = -\Theta$ ; see Restriction 2). Such a test will not only inform us on both the appropriateness of using the discounting model on empirical data, but also say something about the tenability of the “white noise” assumption in dynamical models of affect. Testing this restriction will thus shed light on how to interpret the stochasticity of affect.

Finally, researchers should investigate whether the assumption of exponential decay holds against empirical data (Property 1). For this, researchers should devise models that are build on different discounting functions and test which of these functions fits the data best. As the models in this article are used to investigate the relationship between affect and external stimuli, we can draw inspiration from the delay discounting literature for functional forms that might compete with exponential decay. As an example, we derived the autoregressive counterpart of the quasi-hyperbolic discounting model and suggested it as an alternative for the exponential discounting model that was central to this paper.

It is important to keep in mind that the applicability of each of these dynamical properties may depend on the employed affective measures

and on the studied population. For example, self-reported affective ratings may have a different dynamical structure than psychophysiological measures of affect. Similarly, individuals who suffer from affective disorders might have a different affect dynamical structure than healthy individuals. It is therefore important to test each of the mentioned properties within one’s respective field.

We hope that in our efforts, we have inspired other researchers to investigate how different dynamical models relate to each other. A comprehensive mapping of different computational models, how they are associated, and what restrictions they impose on the dynamical process might hold the key to increasing the adoption of such models in the literature.

### CRediT authorship contribution statement

**Niels Vanhasbroeck:** Conceptualization, Formal analysis, Investigation, Methodology, Visualization, Writing – original draft, Writing – review & editing. **Tim Loossens:** Formal analysis, Investigation, Writing – original draft, Writing – review & editing. **Francis Tuerlinckx:** Conceptualization, Formal analysis, Funding acquisition, Supervision, Writing – original draft, Writing – review & editing.

### Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

### Acknowledgments

The authors would like to thank Sigert Ariens and Kenny Yu for their insightful comments on previous versions of this article.

### Funding

This work was supported by the Research Fund of the KU Leuven under Grant C14/19/054 and by the FWO under Grant G074219N. The funders had no role in study design, data collection, analyses, decision to publish, or preparation of the manuscript.

### Transparency and openness

This study was not preregistered.

### Appendix A. Eq. (1) as a specific case of Eq. (4)

In this section, we show how the general discounting model in Eq. (4) can be rewritten to the specific case in Eq. (1). We start at Eq. (4), which is repeated here:

$$y_t = \alpha + \sum_{j=0}^{t-1} \Gamma^j B x_{t-j} + \epsilon_t .$$

To get to Eq. (1), we first have to define the variables  $y$  and  $x$ , which in this case consist of the happiness score  $h$  as the dependent variable and the independent variables  $cr$ ,  $ev$ , and  $rpe$ . Defining  $y$  and  $x$  allows us to find the dimensions of all parameters in the model. Given that  $d = 1$  and  $k = 3$ ,  $\alpha$  and  $\epsilon$  are of size  $d \times 1 = 1 \times 1$ ,  $\Gamma$  is of size  $d \times d = 1 \times 1$ , and  $B$  is of size  $d \times k = 1 \times 3$ . Filling this out in the above equation, we get:

$$h_t = \alpha + \sum_{j=0}^{t-1} \gamma^j \begin{bmatrix} \beta_1 & \beta_2 & \beta_3 \end{bmatrix} \begin{bmatrix} cr \\ ev \\ rpe \end{bmatrix}_{t-j} + \epsilon_t .$$

Multiplying  $B$  with  $\mathbf{x}$  and splitting the sum leads to:

$$\begin{aligned} h_t &= \alpha + \sum_{j=0}^{t-1} \gamma^j (\beta_1 cr_{t-j} + \beta_2 ev_{t-j} + \beta_3 rpe_{t-j}) + \epsilon_t \\ &= \alpha + \sum_{j=0}^{t-1} \gamma^j \beta_1 cr_{t-j} + \sum_{j=0}^{t-1} \gamma^j \beta_2 ev_{t-j} + \sum_{j=0}^{t-1} \gamma^j \beta_3 rpe_{t-j} + \epsilon_t \\ &= \alpha + \beta_1 \sum_{j=0}^{t-1} \gamma^j cr_{t-j} + \beta_2 \sum_{j=0}^{t-1} \gamma^j ev_{t-j} + \beta_3 \sum_{j=0}^{t-1} \gamma^j rpe_{t-j} + \epsilon_t, \end{aligned}$$

which brings us back to Eq. (1).

### Appendix B. Rewriting the VARMAX in terms of its past

In this section, we will rewrite the VARMAX in terms of its complete history. For clarity and without a loss of generality, we will drop the predictor variables from Eq. (6) and rewrite this model starting from:

$$y_t = \delta + \Theta y_{t-1} + \Phi \epsilon_{t-1} + \epsilon_t.$$

In a first step, we replace  $y_{t-1}$  by its definition, so that:

$$\begin{aligned} y_t &= \delta + \Theta y_{t-1} + \Phi \epsilon_{t-1} + \epsilon_t \\ &= \delta + \Theta (\delta + \Theta y_{t-2} + \Phi \epsilon_{t-2} + \epsilon_{t-1}) + \Phi \epsilon_{t-1} + \epsilon_t \\ &= \delta + \Theta \delta + \Theta^2 y_{t-2} + \Theta \Phi \epsilon_{t-2} + \Theta \epsilon_{t-1} + \Phi \epsilon_{t-1} + \epsilon_t \end{aligned}$$

Doing the same for  $y_{t-2}$ , we get:

$$\begin{aligned} y_t &= \delta + \Theta \delta + \Theta^2 y_{t-2} + \Theta \Phi \epsilon_{t-2} + \Theta \epsilon_{t-1} + \Phi \epsilon_{t-1} + \epsilon_t \\ &= \delta + \Theta \delta + \Theta^2 (\delta + \Theta y_{t-3} + \Phi \epsilon_{t-3} + \epsilon_{t-2}) + \Theta \Phi \epsilon_{t-2} + \Theta \epsilon_{t-1} \\ &\quad + \Phi \epsilon_{t-1} + \epsilon_t \\ &= \delta + \Theta \delta + \Theta^2 \delta + \Theta^3 y_{t-3} + \Theta^2 \Phi \epsilon_{t-3} + \Theta^2 \epsilon_{t-2} + \Theta \Phi \epsilon_{t-2} + \Theta \epsilon_{t-1} \\ &\quad + \Phi \epsilon_{t-1} + \epsilon_t \end{aligned}$$

Repeating this until we reach the initial condition  $y_0$ , this eventually becomes:

$$y_t = \sum_{j=0}^{t-1} \Theta^j \delta + \Theta^t y_0 + \sum_{j=0}^{t-1} \Theta^j \epsilon_{t-j} + \sum_{j=1}^{t-1} \Theta^{j-1} \Phi \epsilon_{t-j}.$$

Note that the index for the last sum starts at  $j = 1$  to accommodate the residuals at time  $t - 1$ .

To simplify this equation, we assume that  $t \rightarrow \infty$ . Together with the restriction that the moduli of the eigenvalues of  $\Theta$  should be smaller than 1, this assumption implies that the first sum converges to  $(I_d - \Theta)^{-1} \delta = \mu$  and that the term  $\Theta^t y_0$  approaches 0 (see Stewart, 1998). Implementing these changes leads to our result:

$$y_t = \mu + \sum_{j=0}^{t-1} \Theta^j \epsilon_{t-j} + \sum_{j=1}^{t-1} \Theta^{j-1} \Phi \epsilon_{t-j}.$$

While we did not consider the predictor variables  $\mathbf{x}$  for simplicity, one can prove that including these predictor variables would lead to the following result:

$$y_t = \mu + \sum_{j=0}^{t-1} \Theta^j \Psi \mathbf{x}_{t-j} + \sum_{j=0}^{t-1} \Theta^j \epsilon_{t-j} + \sum_{j=1}^{t-1} \Theta^{j-1} \Phi \epsilon_{t-j}.$$

### Appendix C. Quasi-hyperbolic discounting model

In this section, we use the approach we used in the article to rewrite a multidimensional quasi-hyperbolic discounting model to an equivalent recursive representation. The quasi-hyperbolic discounting function is defined as (Laibson, 1997):

$$f(t) = \begin{cases} 1 & \text{if } t = 0 \\ N^t K & \text{if } t \in \{1, 2, \dots\} \end{cases},$$

where  $t \in \mathbb{N}$  and where  $K$  and  $N$  are  $d \times d$  matrices that define the shape of the discounting function. Importantly, both  $K$  and  $N$  fall under

the same restrictions as  $\Gamma$  in Eq. (4). For interpretational purposes, we furthermore restrict  $K$  to be a diagonal matrix, so that each of the diagonal elements coincides with the ‘‘recency effect’’ for each variable in  $\mathbf{y}$  (see the paper for a visualization).

For our derivation, it will be useful to rewrite this discounting function to:

$$\begin{aligned} f(t) &= N^t K^{i(t)} \\ i(t) &= \begin{cases} 0 & \text{if } t = 0 \\ 1 & \text{if } t \in \{1, 2, \dots\} \end{cases}, \end{aligned}$$

where the value of  $f(t)$  depends on the indicator function  $i(t)$ , which controls the involvement of  $K$  in the equation. Using this discounting function, we define the quasi-hyperbolic discounting model as:

$$y_t = \alpha + \sum_{j=0}^{t-1} N^j K^{i(j)} B \mathbf{x}_{t-j} + \epsilon_t.$$

To go from this model to its recursive counterpart, we follow the steps of our in-paper derivation. Specifically, we take out the most recent trial from the sum and adjust the index of the sum so that it starts at  $j = 0$ :

$$\begin{aligned} y_t &= \alpha + N^0 K^{i(0)} B \mathbf{x}_t + \sum_{j=1}^{t-1} N^j K^{i(j)} B \mathbf{x}_{t-j} + \epsilon_t \\ &= \alpha + B \mathbf{x}_t + \sum_{j=1}^{t-1} N^j K^{i(j)} B \mathbf{x}_{t-j} + \epsilon_t \\ &= \alpha + B \mathbf{x}_t + \sum_{j=0}^{(t-1)-1} N^{j+1} K^{i(j+1)} B \mathbf{x}_{(t-1)-j} + \epsilon_t. \end{aligned}$$

To get to a recursive model, we have to rewrite this equation in terms of the observation  $y_{t-1}$  at time point  $t - 1$ . For this, the indicator function  $i(j + 1)$  in the sum should become 0 when  $j = 0$ . This is currently not the case, but can be enforced in the following way. First, we extract the values of  $x$  for which  $j = 0$  from the sum:

$$\begin{aligned} y_t &= \alpha + B \mathbf{x}_t + N K^{i(1)} B \mathbf{x}_{t-1} + \sum_{j=1}^{(t-1)-1} N^{j+1} K^{i(j+1)} B \mathbf{x}_{(t-1)-j} + \epsilon_t \\ &= \alpha + B \mathbf{x}_t + N K B \mathbf{x}_{t-1} + \sum_{j=1}^{(t-1)-1} N^{j+1} K^{i(j+1)} B \mathbf{x}_{(t-1)-j} + \epsilon_t. \end{aligned}$$

In the next step, we note that for  $j > 0$ ,  $i(j) = i(j + 1)$ , so that:

$$y_t = \alpha + B \mathbf{x}_t + N K B \mathbf{x}_{t-1} + \sum_{j=1}^{(t-1)-1} N^{j+1} K^{i(j)} B \mathbf{x}_{(t-1)-j} + \epsilon_t.$$

Now, we factor out  $N$  from the sum, which leaves us with:

$$y_t = \alpha + B \mathbf{x}_t + N K B \mathbf{x}_{t-1} + N \sum_{j=1}^{(t-1)-1} N^j K^{i(j)} B \mathbf{x}_{(t-1)-j} + \epsilon_t.$$

Finally, we add some additional terms  $N B \mathbf{x}_{t-1} - N B \mathbf{x}_{t-1}$ ,  $N \alpha - N \alpha$ , and  $N \epsilon_{t-1} - N \epsilon_{t-1}$ , and subsequently factor out  $N$  for the common terms, so that:

$$\begin{aligned} y_t &= \alpha + N \alpha - N \alpha + B \mathbf{x}_t + N B \mathbf{x}_{t-1} - N B \mathbf{x}_{t-1} + N K B \mathbf{x}_{t-1} \\ &\quad + N \sum_{j=1}^{(t-1)-1} N^j K^{i(j)} B \mathbf{x}_{(t-1)-j} + N \epsilon_{t-1} - N \epsilon_{t-1} + \epsilon_t \\ &= \alpha - N \alpha + B \mathbf{x}_t - N B \mathbf{x}_{t-1} + N K B \mathbf{x}_{t-1} \\ &\quad + N \left( \alpha + B \mathbf{x}_{t-1} + \sum_{j=1}^{(t-1)-1} N^j K^{i(j)} B \mathbf{x}_{(t-1)-j} + \epsilon_{t-1} \right) - N \epsilon_{t-1} + \epsilon_t. \end{aligned}$$

Realizing that the terms between the brackets are equal to  $y_{t-1}$ , we get:

$$y_t = \alpha - N \alpha + B \mathbf{x}_t - N B \mathbf{x}_{t-1} + N K B \mathbf{x}_{t-1} + N y_{t-1} - N \epsilon_{t-1} + \epsilon_t.$$

Grouping together the common terms, this equation becomes:

$$y_t = (I_d - N) \alpha + B \mathbf{x}_t + N (K - I_d) B \mathbf{x}_{t-1} + N y_{t-1} - N \epsilon_{t-1} + \epsilon_t,$$

Which is a restricted VARMAX with an additional lag in the predictor variables  $x$ , defined as:

$$y_t = \delta + \Psi x_t + Z x_{t-1} + \Theta y_{t-1} + \Phi \epsilon_{t-1} + \epsilon_{t-1}.$$

For the quasi-hyperbolic discounting model, the parameters are equal to:

$$\delta = (I_d - N)\alpha$$

$$\Psi = B$$

$$Z = N(K - I_d)B$$

$$\Theta = N$$

$$\Phi = -N.$$

## References

- Adolf, J. K., Voelkle, M. C., Brose, A., & Schmiedek, F. (2017). Capturing context-related change in emotional dynamics via fixed moderated time series analysis. *Multivariate Behavioral Research*, 52, 499–531. <http://dx.doi.org/10.1080/00273171.2017.1321978>.
- Ariens, S., Ceulemans, E., & Adolf, J. K. (2020). Time series analysis of intensive longitudinal data in psychosomatic research: A methodological overview. *Journal of Psychosomatic Research*, 137, Article 110191. <http://dx.doi.org/10.1016/j.jpsychores.2020.110191>.
- Asutay, E., & Västfjäll, D. (2022). The continuous and changing impact of affect on risky decision-making. *Scientific Reports*, 12, 10613. <http://dx.doi.org/10.1038/s41598-022-14810-w>.
- Ballard, T., Luckman, A., & Konstantinidis, E. (2023). A systematic investigation into the reliability of inter-temporal choice model parameters. *Psychonomic Bulletin*, 30, 1294–1322. <http://dx.doi.org/10.3758/s13423-022-02241-7>.
- Bennett, D., Davidson, G., & Niv, Y. (2022). A model of mood as integrated advantage. *Psychological Review*, 129, 513–541. <http://dx.doi.org/10.1037/rev0000294>.
- Blain, B., & Rutledge, R. B. (2020). Momentary subjective well-being depends on learning and not reward. *eLife*, 9, 57977. <http://dx.doi.org/10.7554/eLife.57977>.
- Blavatsky, P. R. (2016). A monotone model of intertemporal choice. *Economic Theory*, 62, 785–812. <http://dx.doi.org/10.1007/s00199-015-0931-6>.
- Bonsall, M. B., Wallace-Hadrill, S. M. A., Geddes, J. R., Goodwin, G. M., & Holmes, E. A. (2012). Nonlinear time-series approaches in characterizing mood stability and mood instability in bipolar disorder. *Proceedings of the Royal Society B*, 279, 916–924. <http://dx.doi.org/10.1098/rspb.2011.1246>.
- Booij, S. H. E., S., Jeronimus, B. F., Wichers, M., & Wigman, J. T. (2018). Affective reactivity to daily life stress: Relationship to positive psychotic and depressive symptoms in a general population. *Journal of Affective Disorders*, 225, 474–481. <http://dx.doi.org/10.1016/j.jad.2017.08.051>.
- Cairns, J., & van der Pol, M. (2000). Valuing future private and social benefits: The discounted utility model versus hyperbolic discounting models. *Journal of Economic Psychology*, 21, 191–205.
- Chew, B., Blain, B., Dolan, R. J., & Rutledge, R. B. (2021). A neurocomputational model for intrinsic reward. *The Journal of Neuroscience*, 41, 8963–8971. <http://dx.doi.org/10.1523/JNEUROSCI.0858-20.2021>.
- Congard, A., Dauvier, B., Antoine, P., & Gilles, P.-Y. (2011). Integrating personality, daily life events and emotion: Role of anxiety and positive affect in emotion regulation dynamics. *Journal of Research in Personality*, 45, 372–384. <http://dx.doi.org/10.1016/j.jrp.2011.04.004>.
- de Berker, A. O., Rutledge, R. B., Mathys, C., Marshall, L., Cross, G. F., Dolan, R. J., & Bestmann, S. (2016). Computations of uncertainty mediate acute stress response in humans. *Nature Communications*, 7, 10996. <http://dx.doi.org/10.1038/ncomms10996>.
- de Haan-Rietdijk, S., Voelkle, M. C., Keijsers, L., & Hamaker, E. L. (2017). Discrete- vs. continuous-time modeling of unequally spaced experience sampling method data. *Frontiers in Psychology*, 8, 1849. <http://dx.doi.org/10.3389/fpsyg.2017.01849>.
- Deboeck, P. R., & Bergeman, C. S. (2013). The reservoir model: A differential equation model of psychological regulation. *Psychological Methods*, 18, 237–256. <http://dx.doi.org/10.1037/a0031603>.
- Dejonckheere, E., Houben, M., Schat, E., Ceulemans, E., & Kuppens, P. (2021). The short-term psychological impact of the COVID-19 pandemic in psychiatric patients: Evidence for differential emotion and symptom trajectories in Belgium. *Psychologica Belgica*, 61, 163–172. <http://dx.doi.org/10.5334/pb.1028>.
- Hamaker, E. L. (2012). Why researchers should think “within-person”: A paradigmatic rationale. In *Handbook of research methods for studying daily life* (pp. 43–61). Guilford Press.
- Hamaker, E. L., Dolan, C. V., & Molenaar, P. C. M. (2005). Statistical modeling of the individual: Rationale and application of multivariate stationary time series. *Multivariate Behavioral Research*, 40, 207–233. [http://dx.doi.org/10.1207/s15327906mbr4002\\_3](http://dx.doi.org/10.1207/s15327906mbr4002_3).
- Hamilton, J. D. (1994). *Time series analysis* (2nd ed.). Princeton University Press.
- Keren, H., Zheng, C., Jangraw, D. C., Chang, K., Vitale, A., Rutledge, R. B., Pereira, F., Nielson, D. M., & Stringaris, A. (2021). The temporal representation of experience in subjective mood. *eLife*, 10, Article e62051. <http://dx.doi.org/10.7554/eLife.62051>.
- Koopmans, T. C. (1960). Stationary ordinal utility and impatience. *Econometrica*, 28, 297–309.
- Kuppens, P., Oravecz, Z., & Tuerlinckx, F. (2010). Feelings change: Accounting for individual differences in the temporal dynamics of affect. *Journal of Personality and Social Psychology*, 99, 1042–1060. <http://dx.doi.org/10.1037/a0020962>.
- Kuppens, P., & Verduyn, P. (2015). Looking at emotion regulation through the window of emotion dynamics. *Psychological Inquiry*, 26, 72–79. <http://dx.doi.org/10.1080/1047840X.2015.960505>.
- Laibson, D. (1997). Golden eggs and hyperbolic discounting. *Quarterly Journal of Economics*, 2, 443–477.
- Loossens, T., Dejonckheere, E., Tuerlinckx, F., & Verdonck, S. (2021). Informing VAR(1) with qualitative dynamical features improves predictive accuracy. *Psychological Methods*, 26, 635–659. <http://dx.doi.org/10.1037/met0000401>.
- Loossens, T., Mestdagh, M., Dejonckheere, E., Kuppens, P., Tuerlinckx, F., & Verdonck, S. (2020). The Affective Ising Model: A computational account of human affect dynamics. *PLoS Computational Biology*, 16, Article e1007860. <http://dx.doi.org/10.1371/journal.pcbi.1007860>.
- Lucas, R. E., & Donnellan, M. B. (2007). How stable is happiness? Using the STARTS model to estimate the stability of life satisfaction. *Journal of Research in Personality*, 41, 1091–1098. <http://dx.doi.org/10.1016/j.jrp.2006.11.005>.
- Mazur, J. E. (1987). An adjusting procedure for studying delayed reinforcement. In M. L. Commons, J. E. Mazur, J. A. Nevin, & H. Rachlin (Eds.), *The effect of delay and of intervening events on reinforcement value*. Lawrence Erlbaum.
- Myerson, J., & Green, L. (1995). Discounting of delayed rewards: Models of individual choice. *Journal of Experimental Analysis of Behavior*, 64, 263–276.
- Rutledge, R. B., Moutoussis, M., Smittenaar, P., Zeidman, P., Taylor, T., Hryniewicz, L., Lam, J., Skandali, N., Siegel, J. Z., Ousdal, O. T., Prabhu, G., Dayan, P., Fonagy, P., & Dolan, R. J. (2017). Association of neural and emotional impacts of reward prediction errors with major depression. *JAMA Psychiatry*, 74, 790–797. <http://dx.doi.org/10.1001/jamapsychiatry.2017.1713>.
- Rutledge, R. B., Skandali, N., Dayan, P., & Dolan, R. J. (2014). A computational and neural model of momentary subjective well-being. *Proceedings of the National Academy of Sciences*, 111(33), 12252–12257. <http://dx.doi.org/10.1073/pnas.1407535111>.
- Rutledge, R. B., Skandali, N., Dayan, P., & Dolan, R. J. (2015). Dopaminergic modulation of decision making and subjective well-being. *The Journal of Neuroscience*, 35, 9811–9822. <http://dx.doi.org/10.1523/JNEUROSCI.0702-15.2015>.
- Samuelson, P. A. (1937). A note on measurement of utility. *Review of Economic Studies*, 4, 155–161.
- Scargle, J. D. (1981). Studies in astronomical time series analysis: Modeling random processes in the time domain. *Astrophysical Journal Supplement Series*, 45, 1–71.
- Shumway, R. H., & Stoffer, D. S. (2006). *Time series analysis and its applications* (2nd ed.). Springer.
- Sperry, S. H., Walsh, M. A., & Kwapił, T. R. (2020). Emotion dynamics concurrently and prospectively predict mood psychopathology. *Journal of Affective Disorders*, 261, 67–75. <http://dx.doi.org/10.1016/j.jad.2019.09.076>.
- Steele, J. S., & Ferrer, E. (2011). Latent differential equation modeling of self-regulatory and coregulatory affective processes. *Multivariate Behavioral Research*, 46, 956–984. <http://dx.doi.org/10.1080/00273171.2011.625305>.
- Stewart, G. W. (1998). *Matrix algorithms volume I: Basic decompositions*. Society for Industrial and Applied Mathematics.
- Sutton, R. S., & Barto, A. G. (2018). *Reinforcement learning: An introduction*. MIT Press.
- van den Bos, W., & McClure, S. M. (2013). Towards a general model of temporal discounting. *Journal of Experimental Analysis of Behavior*, 99, 58–73. <http://dx.doi.org/10.1002/jeab.6>.
- Vanhasbroeck, N., Devos, L., Pessers, S., Kuppens, P., Vanpaemel, W., Moors, A., & Tuerlinckx, F. (2021). Testing a computational model of subjective well-being: A preregistered replication of Rutledge et al. (2014). *Cognition and Emotion*, 35, 822–835. <http://dx.doi.org/10.1080/02699931.2021.1891863>.
- Velicer, W. F., & Molenaar, P. C. M. (2012). Time series analysis for psychological research. In *Handbook of psychology* (2nd ed.). American Cancer Society.
- Villano, W. J., Otto, A. R., Ezie, C. E. C., Gillis, R., & Heller, A. S. (2020). Temporal dynamics of real-world emotions are more strongly linked to prediction error than outcome. *Journal of Experimental Psychology: General*, 149, 1755–1766. <http://dx.doi.org/10.1037/xge0000740>.
- Vinckier, F., Rigoux, L., Oudiette, D., & Pessiglione, M. (2018). Neuro-computational account of how mood fluctuations arise and affect decision-making. *Nature Communications*, 9, <http://dx.doi.org/10.1038/s41467-018-03774-z>.
- Voelkle, M. C., & Oud, J. H. L. (2013). Continuous time modelling with individually varying time intervals for oscillating and non-oscillating processes. *British Journal of Mathematical and Statistical Psychology*, 66, 103–126. <http://dx.doi.org/10.1111/j.2044-8317.2012.02043.x>.
- Waugh, C. E., & Kuppens, P. (2021). *Affect dynamics*. Springer Cham, <http://dx.doi.org/10.1007/978-3-030-82965-0>.
- Wendt, L. P., Wright, A. G. C., Pilkonis, W. C., Denissen, J. J. A., Kühnel, A., & Zimmermann, J. (2020). Indicators of affect dynamics: Structure, reliability, and personality correlates. *European Journal of Personality*, 34, 1060–1072. <http://dx.doi.org/10.1002/per.2277>.