Cognition 236 (2023) 105423

Contents lists available at ScienceDirect

Cognition

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

Temporal dynamics of the semantic versus affective representations of valence during reversal learning

Orit Heimer[a,*], Assaf Kron[a,1], Uri Hertz[b,1]

[a] Department of Psychology, University of Haifa, Haifa, Israel
[b] Department of Cognitive Sciences, University of Haifa, Haifa, Israel

ARTICLE INFO

Keywords:
Semantic valence
Affective valence
Reversal learning
Reinforcement learning
Learning rate

ABSTRACT

Valence, the representation of a stimulus in terms of good or bad, plays a central role in models of affect, value-based learning theories, and value-based decision-making models. Previous work used Unconditioned Stimulus (US) to support a theoretical division between two different types of valence representations for a stimulus: the semantic representation of valence, i.e., stored accumulated knowledge about the value of the stimulus, and the affective representation of valence, i.e., the valence of the affective response to this stimulus. The current work extended past research by using a neutral Conditioned Stimulus (CS) in the context of reversal learning, a type of associative learning. The impact of expected uncertainty (the variability of rewards) and unexpected uncertainty (reversal) on the evolving temporal dynamics of the two types of valence representations of the CS was tested in two experiments. Results show that in an environment presenting the two types of uncertainty, the adaptation process (learning rate) of the choices and of the semantic valence representation is slower than the adaptation of the affective valence representation. In contrast, in environments with only unexpected uncertainty (i.e., fixed rewards), there is no difference in the temporal dynamics of the two types of valence representations. Implications for models of affect, value-based learning theories, and value-based decision-making models are discussed.

1. Introduction

Imagine that your neighbors have a new puppy named Max. You usually *experience* pleasure and joy when playing with Max. However, at times when you consider approaching Max, you might experience no pleasure but still *know* that Max is friendly, cute, and you usually feel joy playing with him. That is, you have developed two types of representations of Max being a positive event for you; one as a pleasant feeling and the second as knowledge about potential pleasantness. The representation of a stimulus as positive or (/and) negative is usually termed *valence* (Barrett, 2006b). The above example of Max implies two modes or two types of representations of valence: affective valence (experiencing pleasure playing with Max) and semantic valence (knowing that playing with Max is usually fun) (e.g., Givon, Itzhak-Raz, Karmon-Presser, Danieli, & Meiran, 2019; Itkes, Kimchi, Haj-Ali, Shapiro, & Kron, 2017; Robinson & Clore, 2002b; Russell, 2003; Wang et al., 2021. See Itkes & Kron, 2019 for review.). From an evolutionary perspective, the dissociation between affective and semantic representations is potentially adaptive; it permits considering the values of events without the need to experience a full-blown affective response. For example, people can communicate the value of events with each other or plan future behaviors without the need to activate and experience an affective response.

In psychology and cognitive sciences, there is an accepted assumption that semantic knowledge and affect are distinct categories of the human mind. However, in the affective response, affect and knowledge are often interweaving. For example, processing of an event might involve simultaneous activations of both affect (feelings of fear from X) and semantic (knowing that X is dangerous) representations of its valence. In other words, semantic knowledge might be part of the input that is used in deciding whether to elicit an affective response or not (e.g., Russell, 2003; Sander, Grandjean, Kaiser, Wehrle, & Scherer, 2007, but see Zajonc, 1984). Moreover, some theories assume that conscious experience of emotions represents a conceptual act that, at least partly, involves semantic representations (Barrett, 2006a). The intermix between affect and semantic knowledge in the affective response poses

* Corresponding author at: Department of Psychology, University of Haifa, Abba Khoushy Avenue 199, Haifa 3498838, Israel.
E-mail address: heimer.orit@gmail.com (O. Heimer).

[1] Equal contribution.

https://doi.org/10.1016/j.cognition.2023.105423
Received 19 May 2022; Received in revised form 25 February 2023; Accepted 27 February 2023
Available online 16 March 2023

theoretical and methodological difficulties for the emotion researcher. Theoretically, defining and formulating the differences between affect and knowledge is challenging. Empirically, the distinction between affective and semantic components in measures or tasks is far from obvious.

The current study aims to compare how the two representations of valence are updated, given new information about the stimulus. To illustrate, imagine that your neighbors decided to train Max to be an aggressive guard dog. Such a change in Max's behavior might affect your representations of Max's valence. It can change your feelings toward him (experiencing more unpleasant feelings being with him) and/or your knowledge about him (knowing that Max is no longer cute and that it is not fun to play with him anymore). Using an associative learning paradigm, we systematically examine how variability in rewards (playing with Max generates different degrees of pleasure) and how a reversal of rewards (playing with Max became aversive) affect the temporal dynamics of the affective and semantic representations of valence.

Studying the temporal dynamics of affective and semantic valence representations during associative learning has both a challenge and a promise. The challenge is to tease apart the two aspects of valence that are highly correlated (Itkes et al., 2017; Itkes & Kron, 2019). The promise is that the dissociation provides a new window into studying what is affect, semantic knowledge, the differences between the two, and their distinctive role in learning. In the following, we begin by theoretically distinguishing between affective and semantic representations of valence. We discuss their relation to other theoretical distinctions in the literature and briefly present the empirical challenge of making the dissociation between them and past work that supports this dissociation. Next, we describe the methodological framework of associative learning and discuss the main components of our analytical perspective: sources for prediction error, learning rates, competing conceptual learning frameworks, and reversal learning. We conclude with the present study's outline.

1.1. Semantic versus affective valence representations

This work assumes that affective and semantic valence are different mental representations characterized by specific content and format (Cardinal, Parkinson, Hall, & Everitt, 2002; Gazzaniga, Ivry, & Mangun, 2014).[2] Affective valence is a property of affective response, a profile of time-sensitive physiological and psychological changes after exposure to a stimulus (Itkes et al., 2017). This profile of changes includes autonomic activation such as acceleration or deceleration of the heart rate (e.g., Hodes, Cook III, & Lang, 1985), sweat secretion (Codispoti, Ferrari, & Bradley, 2006; Lempert & Phelps, 2014), hormone secretion (Lewis, 2005), skeletomotor changes as in facial behavior and body posture (e.g., Dael, Mortillaro, & Scherer, 2012; Houtveen, Rietveld, Schoutrop, Spiering, & Brosschot, 2001), and changes in feelings, the conscious experience of affect and emotion. We define the affective valence representation as a non–conceptual property of the transient affective response to a stimulus that indexes this stimulus as positive or negative (Itkes & Kron, 2019).

The semantic representation of valence contains accumulated stored knowledge about the stimulus value (e.g., snakes might be dangerous and, therefore, are negative). Traditionally, accumulated knowledge can be divided into two broad categories: episodic and semantic (Schacter, Wagner, & Buckner, 2000; Tulving, 1983, 1993; Wheeler, Stuss, & Tulving, 1997). Episodic memory is knowledge related to a specific episode at a particular time and place (e.g., I met my friend yesterday at the shopping mall and felt happy to see him). In contrast, semantic knowledge refers to generalized conceptual knowledge about objects and events (e.g., meeting friends can make people feel happy). We argue that it is possible to think and reason about the valence of objects and events without simultaneous full-blown activation of response channels (Itkes & Kron, 2019). People can consistently categorize events according to the valence dimension (whether they are positive or negative) with no change in their experienced feelings, facial expressions, autonomic activation, or other components of the emotional response.

1.1.1. Prior theoretical distinctions

Our proposed working definitions for affective and semantic representations overlap with previous taxonomies. For example, affective versus semantic valence resembles the distinctions between the cold versus hot emotional process (Schaefer et al., 2003), the hot-emotional 'go' system versus the cool-cognitive 'know' system (Metcalfe & Mischel, 1999), self-distance versus self-immersing perspectives (Kross & Ayduk, 2011), impulsive versus reflective systems (Deutsch & Strack, 2004), cognitive appraisal versus feeling (Lazarus & Smith, 1988; Roseman & Smith, 2001), experiential versus non-experiential knowledge (Robinson & Clore, 2002a), and core affect versus affective quality (Russell, 2003, 2005). The conceptualizations most similar to our distinction are those of core affect versus affective quality (Russell, 2003) and experiential versus nonexperiential knowledge (Robinson & Clore, 2002a). However, we use the terms affective and semantic valence representations and not previous taxonomies such as "core affect" and "affective quality" since previous terms are loaded with other assumptions and meanings that may not apply to our purposes (Itkes & Kron, 2019).

Although the abovementioned theories suggest a clear distinction between affective and semantic aspects, this distinction is frequently absent or not well defined in other cases. The term "valence," which is a building block in affective science (Barrett, 2006b) was initially used as semantic valence (for review, see Russell, 1980). For example, Osgood (Osgood, 1952; Osgood, Suci, & Tannenbaum, 1957) referred to valence as a latent dimension of semantic meaning. Later, Russell (1983) used the term valence to indicate the cognitive structure of affect in the language. Only later, the term valence was used as a title for a latent dimension of conscience experience of feelings (Carroll, Yik, Russell, & Barrett, 1999; Russell & Barrett, 1999) and in referring to explicit self-reports on feelings (e.g., Lang, Bradley, & Cuthbert, 1997).

Similar unclarity exists with the term "value," which is therefore considered an "umbrella term" (Ruff & Fehr, 2014). For example, one can use the term "value" to indicate the value of a primary reinforcer like juice upon its consumption (outcome value) (e.g., Kringelbach, O'Doherty, Rolls, & Andrews, 2003). In this case, value refers to the experience, similar to our proposed affective valence. On the other hand, value can also be used to indicate the value of a stimulus in units of a more abstract currency, like money (goal value). It can even include the costs involved in specific options, like delayed consumption (decision value) (e.g., Green & Myerson, 2004). Notably, goal and decision values include an evaluation process beyond the direct experience of consuming the reinforcer. Therefore, they depart from our definition of affective valence representation and involve nonexperiential knowledge that we refer to as semantic valence representation (see Peters & Büchel, 2010 for full details).

[2] We use the term "representation" to refer to a mental object that holds information and has specific content and format (Quilty-Dunn, 2016). Processes are operations done on the representations. Representations can carry information on the characteristics of external objects, i.e., their color, size, and also on their positivity or negativity, i.e., their valence. Importantly, we claim that the valence of a specific external object, i.e., the valence of Max, can be carried in two different mental representations with the same content but in different formats. The first is the affective valence representation, and the other is the semantic valence representation (see Itkes & Kron, 2019; Kron & Weksler, 2022).

1.2. Empirical dissociation between affective and semantic valence representations

Distinguishing between the two types of valence representations becomes even more challenging at the empirical level. It is not always obvious to determine whether a given task or measure is more indicative of affective or semantic valence representation. For example, self-report data on affective feelings often reflect semantic knowledge about valence, not actual feelings (Hamzani, Mazar, Itkes, Petranker, & Kron, 2019; Itkes et al., 2017; Robinson & Clore, 2002b). Another example is tasks that involve cognitive conflict (i.e., congruency effects between response-relevant and response irrelevant features) that are termed "affective" because of the content of the stimuli, e.g., the Affective Simon Task (De Houwer, & Eelen, 1998). Although the Affective Simon task uses stimuli with emotional content, accumulating evidence suggests that the congruency effect reflects semantic conflict (e.g., Duscherer, Holender, & Molenaar, 2008; Itkes et al., 2017).

Recent years produced a growing body of empirical work that provides theoretical and empirical tools to dissociate and distinguish between affective and semantic valence representations. We will shortly review part of it now. Itkes et al. (2017) and Wang et al. (2021) utilized a habituation protocol, i.e., repeated exposure to a stimulus, to tease apart semantic knowledge about the valence of a stimulus and the valence of the emotional response to the same stimulus. Itkes et al. (2017) showed that measures related to affective valence attenuate with repeated exposure to pleasant and unpleasant pictures. In contrast, measures related to semantic valence do not attenuate. In the same vein, Wang et al. (2021) demonstrated dissociation (pre and post-habituation) in the Late Positive Potential (LPP) and activation of brain structures between participants judging the affective valence and participants judging the semantic valence of the same affective pictures.

Returning to the subject of self-reports, which is also relevant to the current study, Hamzani, Mazar, Itkes, Petranker and Kron (2019) compared the associations between self–reports and the physiological response (e.g., heart rate, SCR) to affective stimuli. The self-reports were generated by either feeling–focused instructions (i.e., encouraged participants to report their feelings and not knowledge) or knowledge-focused instructions (i.e., encouraged participants to report semantic knowledge and not feelings). They demonstrated a consistent advantage for feeling-focused over knowledge-focused instructions in predicting the physiological response to affective stimuli. These results strengthen the need to use separate self-reports for the two types of valence, as was done in the current study.

1.3. The current experiments: The temporal dynamics of the semantic versus affective representations of valence during associative learning

The current study examines a possible dissociation in the updating process of the semantic and affective valence representations of a stimulus. Therefore, a prerequisite for our study is a framework in which we can induce controlled and flexible updates in the stimulus' value that can, later on, be analyzed. The habituation protocol (Itkes et al., 2017; Wang et al., 2021) is not suitable for our current purposes because the change in a stimulus' value induced by habituation is limited to attenuation only. However, we can induce and control the updating of a stimulus' value by connecting this stimulus to changing outcomes under our control, i.e., by using the associative learning framework.

During associative learning one stimulus, termed the "conditioned stimulus (CS)," is repeatedly associated with another stimulus, termed the "unconditioned stimulus" (US). The organism learns that the two events, i.e., the CS and the US, are related to one another and therefore, dynamically updates the valence representations of the CS based on the value of the associated US (e.g., Daw & O'Doherty, 2014; Mitchell, De Houwer, & Lovibond, 2009; Rangel, Camerer, & Montague, 2008). Using the associative learning framework, we can systematically manipulate the value of the CS by associating it with varying US values that we fully control. We can then check for dissociations in the updating process of the two valence representations, the affective and the semantic.

Notably, in the current experiments, we used the associative learning framework to induce simultaneous updating of the valence representations of two separate CS ("A" and "B"), each connected to a unique schedule of US. In each trial, the participants chose the CS with the more positive valence representations based on their accumulating experience of the two CS's past contingencies with the varying US values. We will now elaborate on how we model the changes in the valence representations of the CSs within the associative learning framework, namely, the Q-Learning algorithm.

1.3.1. The Q-learning algorithm: Choices, predictions error and learning rate

To model the changes in the valence representations of the CSs, we use a variant of reinforcement learning models (e.g., Rescorla & Wagner, 1972; Sutton & Barto, 2018), namely, the Q-Learning algorithm (Watkins, 1989). A Q learning algorithm consists of three steps: a) Prediction of value, termed *Q-Value*, of all possible CS; b) Selection of the CS that maximizes the predicted reward c) Updating the Q-Values based on experience (Daw & Doya, 2006). Specifically, implementing a Q-Learning algorithm requires three computations: Prediction error, learning rule, and softmax selection rule. We will now explain and formalize each computation, starting with the prediction error.

1.3.1.1. Prediction error. Suppose the participants choose the CS termed "A" in trial t. In this case, the **prediction error** associated with "A" in trial t is given by eq. (1):

$$\delta_t = r_t - Q(A)_t \tag{1}$$

δ_t - Prediction error associated with "A" in trial t.

r_t - Reward in trial t.

$Q(A)_t$ - Predicted Q-Value of "A" in trial t.

The prediction error is the discrepancy between expectations and actual events. Specifically, in our context, the prediction error is the deviation of the actual reward associated with a specific stimulus at a specific time from the predicted Q-Value of this stimulus. The learning process is driven by the prediction error, i.e., by the deviation of the actual rewards from the expected reward; if the reward matches expectation with no deviation, there is no error in prediction and no learning. The actual reward deviates from the expected due to two primary types of uncertainty: expected and unexpected (Soltani & Izquierdo, 2019). *Expected uncertainty* is the uncertainty in rewards attributable to the inherent variability of the given phenomenon. For example, the intensity of pleasure resulting from playing with Max is not fixed and changes from one encounter to the other. Notably, expected uncertainty can be thought of as noise that exists even when the underlying distribution of different outcomes is fixed over time. *Unexpected uncertainty* occurs due to changes in reward probabilities or magnitudes due to changes in the environment, like a reversal of previously learned stimulus-outcome contingencies. For example, the training of Max as a guard dog changed the underlying distribution of the possible intensity of pleasure resulting from playing with him. The experiments in the current study implement *unexpected uncertainty* (reversal) with or without *expected uncertainty* (variability of rewards), as will be detailed in section 1.3.4 below.

1.3.1.2. Learning rule. The Q-Learning algorithm postulates that learning occurs only when the predicted value, $Q(A)_t$, of "A" in trial t is different from the actual reward in trial t, i.e., when the prediction error is not zero. In this case, the predicted Q-Value of "A" should be updated to reflect this prediction error. Eq. (2.a) formalizes the **learning rule,** i.e., the update of the predicted Q-Value of face "A" that is performed in trial t, given the prediction error associated with face "A" in trial t (eq.

1):

$$Q(A)_{t+1} = Q(A)_t + \alpha \cdot \delta_t \tag{2.a}$$

$Q(A)_{t+1}$ - Predicted Q-Value of “A” in trial t + 1.
$Q(A)_t$ - Predicted Q-Value of “A” in trial t.
α - Learning rate; $0 < \alpha < 1$.
δ_t - Prediction error associated with “A” in trial t

On the other hand, because the CS termed “B” was not chosen in trial t, there was no reward and no prediction error connected to it in trial t. Therefore, the learning rule for “B” is given by eq. 2.b, i.e., no change in “B” predicted Q-Value in trial t + 1:

$$Q(B)_{t+1} = Q(B)_t \tag{2.b}$$

$Q(B)_{t+1}$ - Predicted Q-Value of “B” in trial t + 1
$Q(B)_t$ - Predicted Q-Value of “B” in trial t

The decision to what degree to update the stimulus’ Q-Value, given a specific prediction error, reflects an exchange between relying on the long versus short-term rewards history. The learning rate can be considered to control the weight given to the long-term rewards’ history versus the current reward and the resulting prediction error. One extreme option is to completely ignore the prediction error, i.e., to ignore the present reward and base the current Q-Value of the stimulus only on the long-term rewards history. The other extreme option is to sharply update the stimulus’ Q-Value by the total size of the current prediction error. Usually, the learning rate balances the above-mentioned extreme options, i.e., when updating the stimulus’ Q-Value, the prediction error is considered, but not in its full scale (e.g., Gläscher & Büchel, 2005; Niv, 2009). To sum, the learning rate is a crucial concept in the current study because it controls the time it takes for the stimulus’ Q-Value to be updated, i.e., it controls the temporal dynamics of the valence representation during reinforcement learning.

1.3.1.3. Softmax selection rule. Finally, the Q-Learning algorithm can be used to predict the selection of the CS (“A” or “B”). The **softmax selection rule** aims to maximize the predicted reward (eq. 3):

$$P(A)_t = \frac{exp(\beta Q(A)_t)}{exp(\beta Q(A)_t) + exp(\beta Q(B)_t)} \tag{3}$$

$P(A)_t$ - The probability of choosing “A” in trial t
$Q(A)_t$ - Predicted Q-Value of “A” in trial t
$Q(B)_t$ - Predicted Q-Value of “B” in trial t
β - Rate of exploration

Notably, the softmax selection rule estimates the learner’s balance between two complementary goals. The first is the exploitation of previously acquired knowledge of the best action. The second is exploring new actions that might prove beneficial in proportion to their utility. The parameter beta controls the rate of exploration. As beta decreases, selections become more random (i.e., explorative). As beta increases, selections become more deterministic (i.e., exploitative).

1.3.2. Competing conceptual learning frameworks

As elaborated above, our inquiry into the potential dissociation between the temporal dynamics of the affective and sematic valence representations is made using the Q-Learning algorithm, which is based on the critical concepts of Q-Values, prediction error, and learning rate (α). However, the Q-Learning algorithm can be implemented in two competing conceptual learning frameworks: *Common Accumulation* and *Parallel Accumulation*. There is a single Q-Learning algorithm in the *Common Accumulation* framework, whereas three parallel Q-Learning algorithms exist in the *Parallel Accumulation* framework. Fig. 1 contains a mechanistic description of the two alternative frameworks for implementing the Q-Learning algorithm, as will now be elaborated.

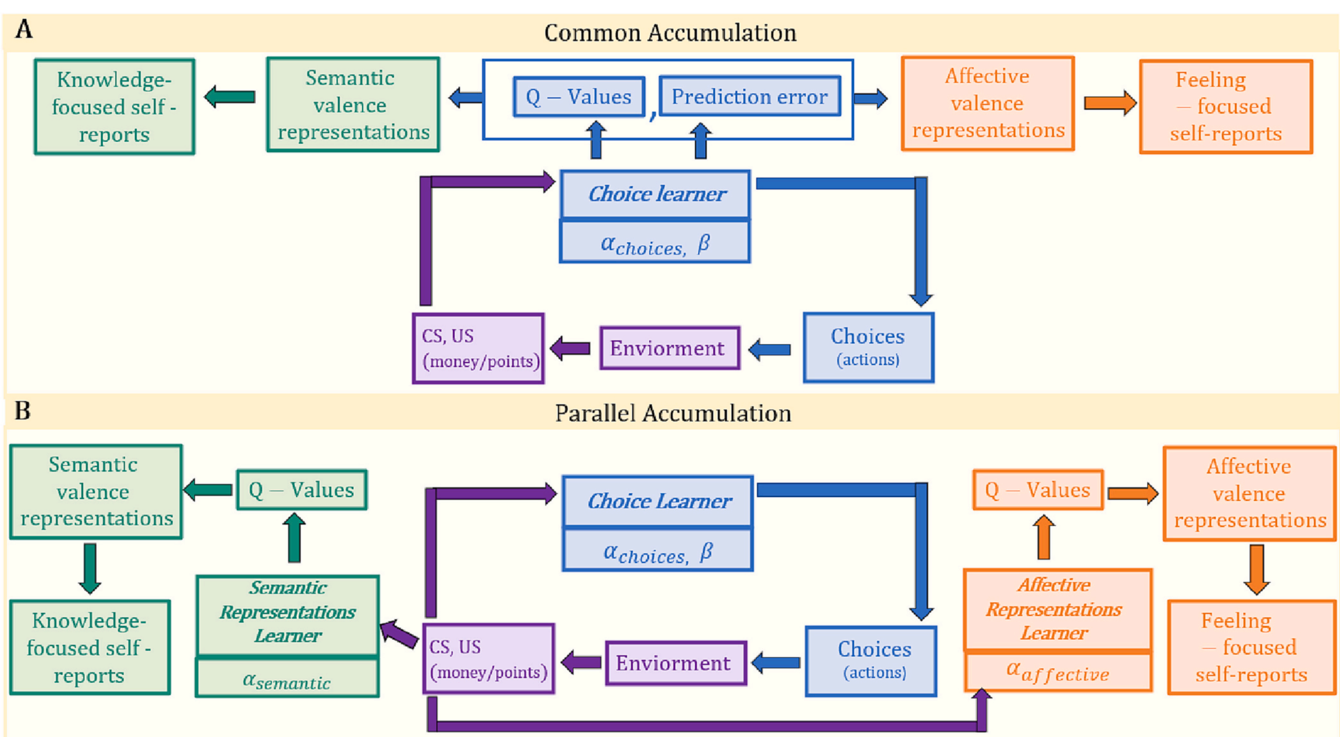

Fig. 1. Alternative conceptual learning frameworks connecting the choices, the affective valence representations, and the semantic valence representations. 1.A. Common Accumulation. According to the *Common Accumulation* framework, there is a single learning process with a single learning rate, which governs the choices. The affective and semantic valence representations are based on components of this single learning process, i.e., both types of valence representations can be based on either the Q -Values or the Prediction error. 1.B. Parallel Accumulation. According to the *Parallel Accumulation* framework, there are three parallel learning processes, one for the choices and one for each type of valence representation. The affective valence representation is based on the Q-Values of the affective learning process, and the semantic valence representation is based on the Q-Values of the semantic learning process.

1.3.2.1. Common accumulation framework. According to the *Common Accumulation* framework, there is a single Q-Learning process, to be termed the *Choice Learner*, in which the Q-Values of the two CSs ("A" and "B") are learned based on trial-by-trial rewards and a single learning rate ($\alpha_{Choices}$). These Q-Values, and the rate of exploration (β) govern the trial-by-trial choices made by the participant. The affective and semantic valence representations are based on components of this single learning process. Notably, this framework is parsimonious because it requires the participants to form only one series of trail-by-trial Q-Values for each CS (See Fig. 1.A).

Past research suggests the semantic valence representations to be based on the Q-Values (Hertz, Bahrami, & Keramati, 2018), whereas the affective valence representations to be based on the prediction error (Rutledge, Skandali, Dayan, & Dolan, 2014). However, it should be stressed that our theoretical distinction between affective and semantic valence representations also allows for other options. We argue that the difference between the affective and semantic valence representation is in their format and not their content. Therefore, we can not preclude in advance reliance of the affective valence representation on Q-Values or reliance of the semantic valence representation on prediction error.

1.3.2.2. Parallel accumulation framework. According to the second learning framework, semantic and affective valence representations are not simply read-out of different latent variables of the *Choice Learner* but are formed directly and independently from the experienced contingencies between the chosen CSs and the USs (points/money in our task). This assumption results in three parallel accumulation, or Q-learning, processes. The first is precisely the same as in the *Common Accumulation* framework, i.e. the *Choice Learner*, that governs the trial-by-trial choices made by the participant. In addition, the *Parallel Accumulation* framework assumes the existence of two additional accumulation processes, one for each type of valence representation, i.e., the *Semantic Representations Learner* and the *Affective Representations Learner*. The semantic valence representations are based on the Q-Values from the *Semantic Representations Learner*, and the affective valence representations are based on the Q-Values from the *Affective Representations Learner*. Notably, the two additional Q-Learning processes rely on the same history of rewards associated with each CS, and the same series of choices. Therefore, the resulting affective and semantic valence representations are not independent. However, the two Q-Learning processes may have different learning rates ($\alpha_{Semantic}$, $\alpha_{Affective}$), i.e., different weighting of short and long-term history of rewards, resulting in different temporal dynamics of the affective and semantic valence representations (See Fig. 1.B).

This framework might seem less likely because it requires more resources to keep track of three series of trail-by-trial Q-values for each CS. However, there is evidence for similar frameworks with multiple representations that keep track of the same underlying process in different timescales (e.g., motor adaptation - Smith, Ghazizadeh, & Shadmehr, 2006; cognitive functions -Soltani, Murray, Seo, & Lee, 2021).

To sum up, according to the *Common Accumulation* framework, the source for possible differences in the temporal dynamics of the two valence representations is their dependency on two different components of the same Q-Learning algorithm, i.e., the *Choice Learner*. On the other hand, according to the *Parallel Accumulation* framework, the temporal dynamics of the two valence representations might differ due to possible different learning rates of the *Semantic Representations Learner* and the *Affective Representations Learner*. We formally compared the two frameworks as part of our analysis (see section 2.1.6 for details).

1.3.3. Reversal learning

The main aim of the current study is to compare the temporal dynamics of affective and semantic representations of valence. To this aim, we implement a specific associative learning paradigm: reversal learning. The reversal learning paradigm (Izquierdo, Brigman, Radke, Rudebeck, & Holmes, 2017) assesses learning flexibility (in our case, learning rate) in the face of *unexpected* uncertainty, i.e., change in a stimulus' value that cannot be attributed to noise. A typical reversal learning session includes an acquisition phase immediately followed by a reversal phase. Two (or more) stimuli are presented during the acquisition phase, each connected to a different schedule of Unconditioned Stimulus (US) (e.g., Atlas, Doll, Li, Daw, & Phelps, 2016; M. R. Delgado, Labouliere, & Phelps, 2006; Mertens & De Houwer, 2016; Schiller, Levy, Niv, LeDoux, & Phelps, 2008). During the acquisition phase, the participants learn the value (valence) of each CS. For example, in fear reversal learning, there are two Conditioned Stimuli (CS). One CS is always (or partly) connected to an aversive stimulus such as an electric shock, and therefore it is marked with a "+" sign (CS+). The second CS is never connected to an aversive stimulus, and therefore it is marked with a "-" sign (CS-).

The acquisition phase is followed by a (usually uninformed) reversal phase. During the reversal phase, the previously learned schedule of US connected to each CS is reversed. Each CS is now connected to the alternate schedule of the US. Continuing the fear reversal learning example, the former CS+ (new CS-) is no longer paired with the aversive US in the reversal phase. Instead, the previous CS- (new CS+) is now paired with the aversive US (e.g., Atlas, 2019; Costa, Bradley, & Lang, 2015; Schiller & Delgado, 2010). Notably, the inherent complexity of the reversal learning paradigm, i.e., the need to adjust the valence of the two CS stimulatingly, is ideal for comparing the two types of valence learning rates.

1.3.4. The current experiments

Two experiments systematically investigated the learning rates of the affective and semantic representations of valence during reversal learning. Experiment 1 compared the learning rates in an experimental environment that involves both expected (i.e., variability of the monetary rewards) and unexpected (i.e., reversal) uncertainty. Experiment 2 examined the potential moderation of expected uncertainty on the effect of unexpected uncertainty (reversal) on the learning rates. Preuschoff and Bossaerts (2007) and Diederen and Schultz (2015) showed that in cases involving expected variability, it is advantageous to adjust the prediction error based on the expected reward variability. The higher the variability, the smaller the prediction error and vice versa. This adjustment is computationally translated to a lower rating rate, the higher the variability of the rewards. Experiment 2 enabled us to check for the existence of this adjustment in the learning rates of the semantic and affective valence representations.

2. Experiment 1

In experiment 1, we examined the effect of unexpected uncertainty (reversal) and expected uncertainty (variability of rewards) on the learning processes of the two types of valence representations for the CS. In this experiment, the participants made 80 choices of one out of two neutral faces of women (face "A" and face "B," the CSs). Each face was connected to a unique schedule of variable monetary rewards (the US), unknown to the participants, such as one face was more profitable than the other. After 40 trials, the contingencies of the monetary rewards to the faces were reversed without informing the participants, turning the second face into the profitable choice. Notably, in each trial, after choosing the face and receiving the monetary reward the participants made feeling-focused (i.e., affective valence) and knowledge-focused (i.e., semantic valence) self–reports on the valence of the chosen face.

2.1. Method and material

2.1.1. Participants

Forty-four students from the University of Haifa [13 male, aged 26.2 ± 5.45 (mean ± SD); 33 female, aged 21.6 ± 2.3] participated in the first experiment. The study was approved by the University of Haifa's, Faculty of Social Sciences Research Ethics Committee (Project ID Number: 318/18). All participants were screened for neurological disorders. Participants received course credit or were paid around NIS 40 (~$12) for their participation. In addition, the participants received a bonus of up to NIS 30 (~$9), depending on the rewards in two randomly selected trials. We excluded nine participants from the study due to lack of learning[3] or lack of variance in their self-report ratings.[4] The final sample size for the analysis was $n = 35$. Because we are interested in the possible differences between the affective and semantic valence learning rates, i.e., a within-subject design, and assuming a power of 80%, a sample size of n = 35 should capture a medium effect size, i.e., Choen's d of around 0.5 (Erdfelder, FAul, Buchner, & Lang, 2009).

2.1.2. Conditioned stimuli (CSs)

The CSs were two neutral faces of women (face "A" and face "B"). We randomly selected a unique pair of faces to be displayed to each participant out of a pool of four women's faces, all taken from the Karolinska Directed Emotional Faces (KDEF) database (Lundqvist, Flykt, & Öhman, 1998). We used the "without hairline" version of the KDEF faces, validated in Goeleven et al.'s study (Goeleven, De Raedt, Leyman, & Verschuere, 2008). Also, the selected faces received very similar ratings in several criteria, including attractiveness, emotional intensity, and valence, in Garrido and Parda's study on KDEF (Garrido & Prada, 2017). To verify that the participants perceived the two faces we randomly chose for them (face "A" and face "B") as equally neutral at the beginning of the experiment, we created a database composed of all the initial self–ratings the participants gave in both experiments. We then ran two ANOVAs, one for the semantic valence initial self-reports and one for the affective valence initial self-reports, and checked for an effect of face identity on the ratings. In both ANOVAs, the effect of the face identity on the ratings was insignificant (*p*-values of 0.64 and 0.29 for the semantic and affective valence, respectively, see Supplementary Table 1).

2.1.3. Feeling-focused self-reports

Based on Itkes et al. (2017), the participants rated the affective valence of the CSs using two unipolar scales, one for positive feelings and the second for negative feelings. Both scales have the appearance of a volume graph ranging from low to high (coded 0-no feelings to 8-high, see Supplementary Fig. 1). Because the CS was a woman's face, the participants were instructed to differentiate between the feelings the woman in the image might have and their feelings while looking at the woman, and to report only their feelings. On the positive scale, the participants were instructed to rate whether they felt feelings of pleasure, happiness, or any other pleasant feeling while looking at the CS. If they did not feel anything pleasant, or were unsure if they felt something pleasant, they were instructed to report "0". Only if they were sure they felt feelings of pleasure, happiness, or any other pleasant feeling while looking at the CS, they were instructed to rate the experienced intensity of their positive feelings from 1 (low) to 8 (high). On the negative scale, the participants were instructed to rate whether they felt feelings of displeasure, sadness, or any other unpleasant feelings while looking at the CS and if so, to rate the experienced intensity of their negative feelings. We calculated the affective valence of the CS as the difference between the positive and negative ratings, resulting in a measure ranging from −8 to 8. (For a similar transformation, see Haj-Ali, Anderson, & Kron, 2020; Larsen, Norris, & Cacioppo, 2003; Kron, Goldstein, Lee, Gardhouse, & Anderson, 2013).

Importantly, before reporting their feelings on the positive and negative feeling scales, the participants were presented with a scale of "general" feelings. In this scale, the participants were instructed to indicate whether they had any feelings while looking at the CS and, if so, to rate the experienced intensity of their feelings. This scale makes the participants contemplate whether they indeed felt something before rating their positive and negative feelings. Evidence for the validity of this set of scales as a measure of affective valence can be found in Itkes et al. (2017).

2.1.4. Knowledge-focused self-reports

We measured semantic evaluations by requesting the participants to report their expectations regarding the connections between the CS and the US, coupled with confidence in this expectation. Specifically, we asked them to evaluate what they thought would happen, further selecting a specific alternative (CS). The options ranged from "I will certainly lose money" (coded as 0), through the option "I do not know what will happen" (coded as 4), to the option "I will certainly receive money" (coded as 8). (See Supplementary Fig. 2 for a similar scale used in experiment 2). A comprehensive analysis of this measure's validity can be found in Boddez et al. (2013). Additional support for this measure can be found in Hertz et al. (2018).

2.1.5. The reversal-learning task

The reversal task consists of two phases of 40 trials each: an acquisition phase followed by a reversal phase. During the whole task, the participants made 80 sequential choices of one of two neutral faces of women, face "A" and face "B," which served as the Conditioned Stimuli (CS) (see Fig. 2). We paired each face with a different pseudo-random schedule of monetary reinforcements, unknown to the participants, which served as the Unconditioned Stimuli (US).[5] During the first 40 trials that composed the acquisition phase, choosing face "A" yielded an average profit of NIS 10 (~ $3). Crucially, the actual amounts varied uniformly between NIS 5 and 15 (SD ~ 3). Selecting face "B" yielded an average break-even (zero) result. Again, exact amounts varied uniformly between NIS -5 and 5 (SD ~ 3). During the last 40 trials, the reversal phase, we reversed these contingencies without informing the participants. In 60% of the 80 trials, the participants could choose one of the two faces (i.e., free–choice trial, instrumental conditioning). In the rest of the trials, the participants were obligated to select one of the faces (i.e., forced-choice trial, classical conditioning). Before their first choice, the participants made an initial self-rating of faces "A" and "B"'s semantic and affective valence to ensure that the participants indeed perceived the two faces to be used as the CS (i.e., face "A" and face "B") as equally neutral. The internal order of the self-report ratings was counterbalanced.[6] (See Supplementary Table 2 for a summary of the features of experiment 1 and experiment 2.)

The free-choice trials started with the presentation of the two faces (face "A" and face "B") on a horizontal line in the middle of the screen. We randomly selected the position of the faces (the left side or the right side) in each trial. After the participant made her choice, The chosen CS alone (4 s.), the chosen CS and the US together (2 s.), and the chosen CS alone (5 s.) were presented. Then, the participants made self–reports on the valence of the chosen face. (See Fig. 2 for a similar sequence used in experiment 2).

The forced-choice trials occurred in the second and fourth trial of each block. They had the same sequence as the free-choice trial described above, with one difference. One of the faces was marked with a big "x" on the choosing screen, indicating that the participant was not

[3] The percentage of free choices of the profitable face (i.e., face "A") during the second half of the acquisition phase was below chance level (50%).

[4] Defining 16 clusters of self-reports (4 scales [semantic, feelings, positive, negative] * 2 phases [Acquisition, Reversal] * 2 CS [Face A, Face B]). Calculating the variance in each of the defined clusters for each participant. Excluding participants with zero variance in >20% of the clusters.

[5] We used two different pseudo-random schedules for each face to avoid the possibility that the exact order of reinforcements could explain the results.

[6] We used four different versions (a - Semantic, Feelings, Positive, Negative, b - Semantic, Feelings, Negative, Positive, c - Feelings, Positive, Negative, Semantic, d - Feelings, Negative, Positive, Semantic).

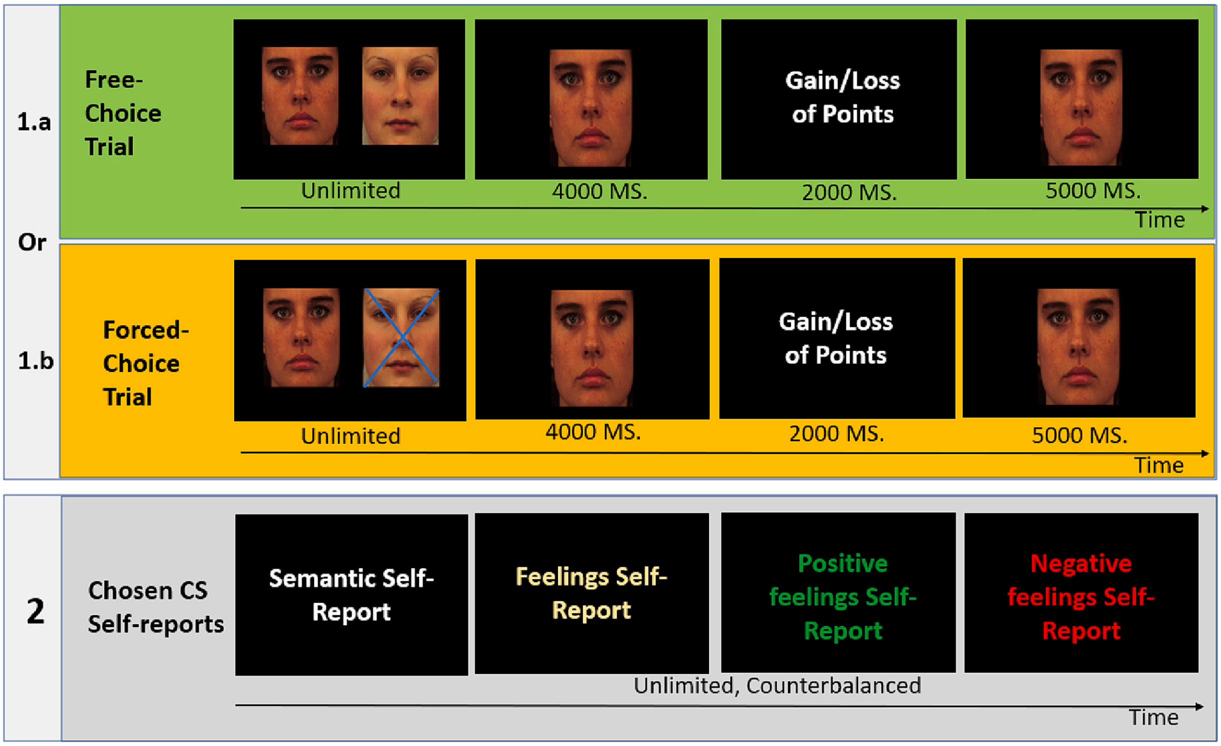

Fig. 2. Trial sequence. 2.a. The participants choose between two neutral women's faces, face "A" and face "B." Next, the selected face appeared for 4 s., followed by a presentation of the reward (gain or loss of points) for 2 s. The selected face appears again for 5 s. 2.b. The same sequence as in 1.a., but one of the faces is marked with X and the participant cannot choose it. 2.c. The participants perform the two types of valence self–reports, semantic (knowledge-focused) and affective (feeling-focused), regarding the chosen face, with no time limit.

allowed to choose it, leaving her with only one face out of two to "choose." In each block, one of the forced-choice trials forced the participant to choose face "A," and the other forced her to choose face "B."

2.1.6. Analysis strategy

Our primary research question investigated the temporal dynamics of the two representations of valence. To meet this aim, we first performed an illustrative comparison of the reliance of the two types of valence on the reward history by cross-correlation analysis. We then implemented three Q-Learning processes for each participant's trial–by–trial data using the Q-Learning algorithm (Watkins, 1989): the *Choice Learner,* the *Semantic Representations Learner* and the *Affective Representations Learner*. Next, we compared two competing alternatives for a conceptual framework of learning. According to the first alternative, the *Common Accumulation* framework, there is a single Q-Learning process, the *Choice Learner,* with a single learning rate (Fig. 1.A). The affective and semantic valence representations are based on components of this single Q-Learning process. According to the second alternative, the *Parallel Accumulation* framework, there are, in addition to the *Choice Learner*, two separate Q-Learning processes, the *Semantic Representations Learner* and the *Affective Representations Learner,* one for each type of valence representation (Fig. 1.B). Notably, only this framework allows for different learning rates for each type of valence representation. Lastly, we compared the learning rates from the *Semantic Representations Learner* and the *Affective Representations Learner* under the framework that received more support from our data, i.e., the *Parallel Accumulation* framework.

2.1.6.1. Cross-correlation analysis. We first performed an illustrative comparison of the reliance of the two types of valence on reward history using cross-correlation analysis. The cross-correlation analysis is based on the notion that the learning rate controls the weight given to the history of rewards versus the current trial's reward in determining the current valence of the CS. The lower the learning rate, the more weight given to the reward history, i.e., we expect a higher correlation between the current trials' valence and past rewards (USs). On the other hand, the higher the learning rate, the less weight given to the reward history, i.e., we expect a lower correlation between the current trials' valence and past rewards (USs). Based on these considerations, we checked for a correlation between the self-report of the current trial's valence and the current trial's reward (0–lag cross-correlation). The correlation calculation was done on the individual participant's level ($N = 35$), resulting in a distribution of 35 correlations, each based on the correlation between 80 ratings and rewards connected to a specific participant. We then conducted a Fisher's Z transformation on the correlations. We did the same procedure to check the correlation between the self-report of the current trial's valence and the reward in different lags, e.g., the last trial where this option was chosen (1–lag cross-correlation). We then formally compared the distributions of two extremes, i.e., the 0–lag and 3–lag cross-correlations (after Fisher's Z transformation), for each valence type, using paired *t*-tests. The lower the learning rate, the more we expect the 0–lag and 3–lag cross-correlations to be the same. On the other hand, the higher the learning rate, the more we expect the 0–lag cross-correlation to be higher than the 3-lag cross-correlation.

2.1.6.2. Computational modeling. We implemented three Q-Learning processes for each participant's trial–by–trial data: the *Choice Learner*, the *Semantic Representations Learner* and the *Affective Representations Learner*. The Computational modeling was done via a Q-Learning reinforcement learning algorithm (Watkins, 1989). We, therefore, implemented computations of the prediction error (eq. 1) and the learning rule (eqs. 2.a and 2.b) detailed in section 1.3.1. For the *Choice Learner* the algorithm also included the softmax selection rule (eq. 3). We will now complete our description of the computational modeling by detailing the different optimization and parameter estimation processes used in each Q-Learning process.

2.1.6.2.1. The choice learner. For the *Choice Learner* (Fig. 1.A), we used each participant's trial-by-trial choices and rewards data and the Q-Learning algorithm (i.e., eqs. 1 to 3) to fit a Q-Learning process to the choices made by each participant. In this learning process, the fitting was done by optimizing the log-likelihood of the probability of choices, given by eq. 3. The resulting parameters from this Q-Learning process are alpha and beta for each participant. In addition, the *Choice Learner* provides estimated trial-by-trial Q-Values and prediction errors for each participant, which could be later used to directly compare the two competing learning frameworks (see section 2.1.6.2.4).

2.1.6.2.2. The semantic/affective representations learner. For the *Semantic Representations Learner* and the *Affective Representations Learner* (Fig. 1.B), we used a variant of the Q-Learning algorithm (detailed in section 1.3.1) that was fitted to the knowledge-focused and feeling-focused self-reports instead of the choices. Therefore, this variant included only the prediction error (eq. 1) and the learning rule (eq. 2.a + 2.b) and did not include the softmax decision rule (eq. 3) and estimation of the rate of exploration (β). In addition, this variant used a different error function for the parameter estimation. More specifically, fitting the *Semantic Representations Learner* and the *Affective Representations Learner* to the knowledge-focused and feeling-focused self-reports data was performed by minimizing the **square error** of distance between the self-ratings data and the Q-Values derived from the algorithms (a different set of Q-Values for each Q-Learning process).

The *Semantic Representations Learner* **square error** is given in eq. (5):

$$error_{semantic} = error_{semantic}(A) + error_{semantic}(B) \tag{5}$$

$error_{semantic}$ – The total *Semantic Representations Learner* square error, including all the trials.

$error_{semantic}(A)$ – The *Semantic Representations Learner* square error, relevant **only** for trials where face "A" was chosen and rated. See details in eq. 5.1

$error_{semantic}(B)$ - The *Semantic Representations Learner* square error, relevant **only** for trials where face "B" was chosen and rated. See details in eq. 5.2.

$$error_{semantic}(A) = \sqrt{\sum_{c=A} \left(Rate_{semantic}(c) - (b_0 + b_1 \bullet Q_{semantic}(c)\,)^2 \right.} \tag{5.1}$$

$Rate_{semantic}(c)$ – knowledge-focused self-reports of face "A" in trial t

b_0, b_1- Scaling of the Q-Values to match the self–reports units of measurement

$Q_{semantic}(c)$ - Predicted Q-Value of face "A" in trial t

$$error_{semantic}(B) = \sqrt{\sum_{c=B} \left(Rate_{semantic}(c) - (b_0 + b_1 \bullet Q_{semantic}(c)\,)^2 \right.} \tag{5.2}$$

$Rate_{semantic}(c)$ – knowledge-focused self-reports of face "B" in trial t

b_0, b_1- Scaling of the Q-Values to match the self–reports units of measurement

$Q_{semantic}(c)$ - Predicted Q-Value of face "B" in trial t

We used a similar set of equations to fit the *Affectitive Representations Learner* Q-Learning process:

$$error_{affective} = error_{affective}(A) + error_{affective}(B) \tag{6}$$

$error_{affective}$ – The total *Affective Representations Learner* square error, including all the trials.

$error_{affective}(A)$ – The *Affective Representations Learner* square error, relevant **only** for trials where face "A" was chosen and rated. See details in eq. 6.1

$error_{affective}(B)$ - The *Affective Representations Learner* square error, relevant **only** for trials where face "B" was chosen and rated. See details in eq. 6.2.

$$error_{affective}(A) = \sqrt{\sum_{c=A} \left(Rate_{affective}(c) - \left(b_0 + b_1 \bullet Q_{affective}(c)\,\right)^2 \right.} \tag{6.1}$$

$Rate_{affective}(c)$ – feeling -focused self-report of face "A" in trial t

b_0, b_1- Scaling of the Q-Values to match the self–report units of measurement

$Q_{affective}(c)$ - Predicted Q-Value of face "A" in trial t

$$error_{affective}(B) = \sqrt{\sum_{c=B} \left(Rate_{affective}(c) - \left(b_0 + b_1 \bullet Q_{affective}(c)\,\right)^2 \right.} \tag{6.2}$$

$Rate_{affective}(c)$ – feeling -focused self-report of face "B" in trial t

b_0, b_1- Scaling of the Q-Values to match the self–report units of measurement

$Q_{affective}(c)$ - Predicted Q-Value of face "B" in trial t

The resulting parameters from each Q-Learning process are alpha, $b_{0,}$ and b_1 for each participant. In addition, these Q-Learning processes result in estimated trial-by-trial Q-Values and prediction errors for each participant, which could be later used to directly compare the two competing learning frameworks (see section 2.1.6.2.4).

2.1.6.2.3. Pre-processing. We used self-rating Z scores in eqs. (5) and (6). As we fitted the Q-Learning algorithms for each participant separately, we calculated the Z scores based on all the knowledge-focused (or feeling-focused) self-reports given by this specific participant to both faces (i.e., 80 trials = 80 observations for each valence type).

2.1.6.2.4. Frameworks comparison. To decide which of the two conceptual learning frameworks, the *Common Accumulation* framework or the *Parallel Accumulation* framework, best fit our data, we systematically run a series of mixed effects regression models with the affective (/semantic) valence self-reports as the dependent variable. We used group-level coefficients (fixed effects) to model population-level effects and individual-level coefficients (random effects) to capture average individual responses (Gelman & Hill, 2006). We report standardized coefficients, which represent the partial correlation between the dependent and independent variables and are, therefore, indicators of effect size. We compared the model fitting scores BIC and AIC between the models using ANOVA (BIC-Bayesian information criterion, Schwarz, 1978; AIC - Akaike information criterion, Akaike, 1974).

In the case of explaining the feeling-focused self-reports, the competing candidates for independent variables in the mixed effects regression models were: 1) the Q-Values from the *Choice Learner* process, 2) the prediction errors from the *Choice Learner* process, 3) the Q-Values from the *Affective Representations Learner* process, 4) the prediction errors from the *Affective Representations Learner* process. Notably, all the above variables were used as both fixed and random effects. Suppose the *Common Accumulation* framework best fits our data. In that case, we expect the best regression model will include either the prediction errors or the Q-Values from the only Q-Learning process under this framework, i.e., the *Choice Learner* process (Fig. 1.A). If, on the other hand, the *Parallel Accumulation* framework best fits our data, we expect the best mixed regression model will include the Q-Values from the *Affective Representations Learner* process (Fig. 1.B).

We repeated the process when explaining the knowledge-focused self-reports. We used the same variables from the *Choice Learner* process mentioned above. However, this time we replaced the variables from the *Affective Representations Learner* process with their parallels in the *Semantic Representations Learner* process. Our prediction for the best variable under the *Common Accumulation* framework is either the Q-Values or the prediction errors from the only Q-Learning process under this framework, i.e., the *Choice Learner* process. If, on the other hand, the *Parallel Accumulation* framework best fits our data, we expect the best regression model will include the Q-Values from the *Semantic Representations Learner* process.

2.2. Results

2.2.1. Exclusion criteria

We commenced the analysis by screening the data and excluding seven participants who failed to show a sufficient learning pattern, i.e., their percentage of free choices of the profitable face (face "A") during the second half of the acquisition phase was below the chance level (50%). Next, we checked the quality of the self-reports given by the participants. Based on the dynamic nature of our task, we expected to find variability in the self-reports related to each face in each phase. We defined 16 clusters of self-reports (4 scales [semantic, feelings, positive, negative] * 2 phases [Acquisition, Reversal] * 2 CS [Face A, Face B]), and calculated the variance in each of the defined clusters for each participant. We excluded 2 participants who exhibited low variability, i.e., zero variance in >20% of the clusters. Altogether, we excluded 9 participants, i.e., we were left with a final sample of 35 participants. Notably, we repeated the primary analysis presented below of comparing the learning rates of the two types of valence **without exclusion** of any participants and received similar results (See supplementary file section C for details).

2.2.2. Block-to-block trends of the main variables

As can be seen in Fig. 3.A, on average, the participants learned to choose the preferred option in each phase (option A in the acquisition phase and option B in the reversal phase), as seen by the choices trends. In addition, the participants' self-reports of the preferred option in each phase were more positive than their self-reports for the unpreferred option. The affective valence self-reports were updated faster than the semantic ones during the acquisition phase and after the reversal.

2.2.3. Cross-correlation analysis

The learning rate controls the correlation of past rewards with the current valence ratings. The higher the learning rate, the less correlation past rewards have with the current trial's valence self-report. To illustrate the difference between the reliance of the two types of valence on reward history, we computed correlations between the current trial's valence self-report and the history of rewards (USs) for the two types of valence. The correlation calculation was done on the individual participant's level ($N = 35$) (see analysis strategy, section 2.1.6.1 for more details). As illustrated in Fig. 3.B, the correlation between the feeling-focused self-reports and past rewards weakens as the gap between the self-report and the time the reward has been presented increases. A comparison between the two extremes (0-lag cross-correlation and 3-lag cross-correlation) shows a significant drop between the 0-lag cross-correlation ($M = 0.60$, $SD = 0.38$) and the 3-lag cross-correlation ($M = 0.45$, $SD = 0.32$) ($t(33) = 4.27$, $p < .001$, $d = 0.73$).[7] In contrast, the correlation between the knowledge-focused self-report and past rewards remains constant as the gap between the self-reports and the time the US has been presented increases (0-lag: $M = 0.68$, $SD = 0.24$, 3-lag: $M = 0.69$, $SD = 0.22$, $t(34) = -0.60$, $p = .55$, $d = -0.1$).

2.2.4. Computational modeling

2.2.4.1. Frameworks comparison. We fitted the three Q-Learning processes (i.e., the *Choice Learner*, the *Affective Representations Learner,* and the *Semantic Representations Learner*) to the trial-by-trial choices and ratings of the 35 participants. The fitting procedure, detailed in section 2.1.6.2, resulted in estimated parameters (i.e., learning rates for each Q-Learning process, rate of exploration ($\boldsymbol{\beta}$), and scaling parameters ($\boldsymbol{b_0}$, $\boldsymbol{b_1}$)) for each participant. We also recovered the trial-by-trial Q-Values and prediction errors for each learning process and used them for the frameworks comparison, as detailed below.

Table 1 summarizes the results of eight mixed effects regression models with the affective (/semantic) valence self-reports ratings as the dependent variable, as detailed in section 2.1.6.2.4. The best single variable to explain the feeling–focused self-reports is our prediction for the best fit under the *Parallel Accumulation* framework, i.e., the Q-Values from the *Affective Representations Learner* process. Under the *Common Accumulation* framework, the best-explaining variable is the Q -Values of the *Choice Learner* process. Notably, the gap in AIC/BIC terms between the Q -Values from the *Affective Representations Learner* process and the Q -Values from the *Choice Learner* process is only 2%. This difference can be attributed to the fact that the *Affective Representations Learner* was directly fitted to the feeling–focused self-reports. In contrast, the *Choice Learn*er was fitted to the choices. For full details of the models, see supplementary tables D.1 to D.4.

The best single variable to explain the knowledge–focused self-reports is again our prediction for the best fit under the *Parallel Accumulation* framework, i.e., the Q-Values from the *Semantic Representations Learner* process. Under the *Common Accumulation* framework, the best-explaining variable is again the Q-Values from the *Choice Learner* process. The gap in AIC/BIC terms between the Q -Values from the *Semantic Representations Learner* process and the Q -Values from the *Choice Learner* is 4%. As in the case of the feeling-focused self-reports, this difference can be attributed to the fact that the *Semantic Representations Learner* was directly fitted to the knowledge–focused self-reports. For full details of the models, see supplementary tables D.5 to D.8.

In summary, the differences between the frameworks are minor and can be attributed to the direct fit of the self-reports to the learning processes in the *Parallel Accumulation* framework. However, it should be noted that the *Common Accumulation* framework indicated that both types of self-reports are based on the same component of the *Choice Learner* process, i.e., its Q-Values, and therefore should be the same. This result contradicts our data that shows differences between the knowledge–focused and feeling–focused self–reports in both the block-to-block trends and the cross-correlation analysis. Therefore, we concluded that the *Parallel Accumulation* framework received more support from our data and chose it as the framework we adopted for the learning rates comparison detailed below.

2.2.4.2. Comparing the learning rates of the semantic and affective valence. We moved to check whether the learning rate of the affective valence differs from the learning rate of the semantic valence under the framework of *Parallel Accumulation*. To this end, we compared the estimated alpha parameters (i.e., the learning rate) derived from the *Affective Representations Learner* Q-Learning process and the *Semantic Representations Learner* Q-Learning process. We found that the estimated alpha from *Affective Representations Learner* ($M = 0.48$, $SD = 0.32$) is higher than the estimated alpha from the *Semantic Representations Learner* ($M = 0.28$, $SD = 0.19$), and this difference is statistically significant ($t(34) = -3.39$, $p = .002$, $d = -0.57$). (see Fig. 3.C and supporting bayesian inference in supplementary fig. 4).[8,9] We also conducted this analysis without excluding participants and found similar effects. See Supplementary file section C for details.

[7] One of the participants was omitted because she has zero variance in the feeling-focused self- reports.

[8] For additional information on the results of the Q-Learning processes see supplementary table 3.

[9] We have checked the possibility that a Q-Learnig process with two learning rates, one connected to positive prediction error and one to negative prediction error, is a better fit for our data. We used the BIC criteria (Schwarz, 1978). The results show that a Q-Learning process with two learning rates does not over-perform a Q-Learning process with one learning rate in both experiments. Therefore, we did not adopt this alternative. For the analysis, see supplementary table 4.

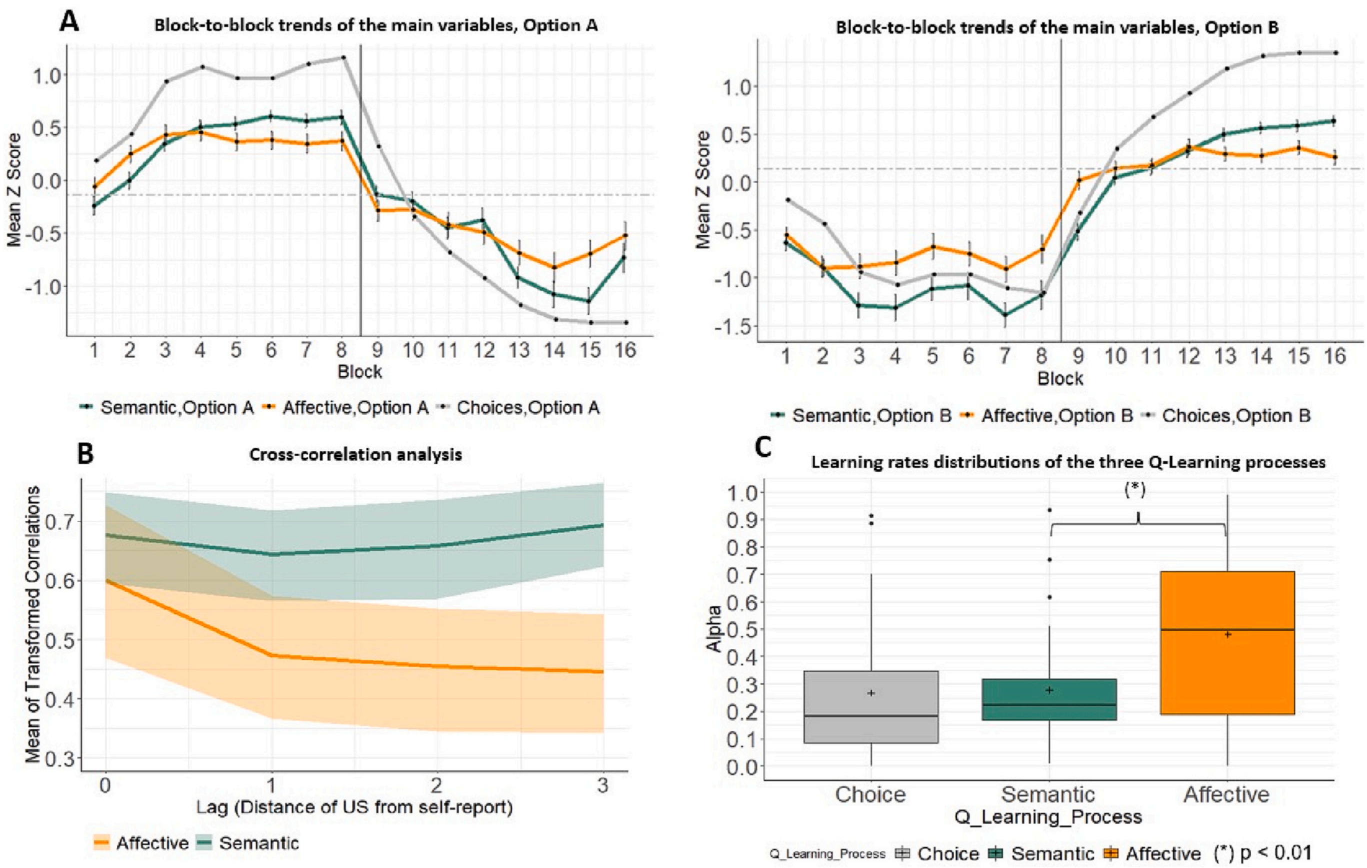

Fig. 3. Experiment 1 results. 3.A. Block-to-block trends of the main variables. On average, the participants learned to choose the preferred option in each phase (option A in the acquisition phase and option B in the reversal phase), as seen by the choices trends (grey line). The dotted horizontal grey line indicates the chance level (50%). In addition, the participants' self-reports of the preferred option in each phase were more positive than their self-reports for the unpreferred option. The affective valence self-reports were updated faster than the semantic ones during the acquisition phase and after the reversal. 3.B. The correlation between the affective valence self-report and past rewards (i.e., USs) weakens as the gap between the self-report and the time the US has been presented increases. In contrast, the correlation between the semantic valence self-report and past rewards remains constant as the gap between the self-reports and the time the US has been presented increases. 3.C. the learning rate (alpha) of the Affective Representations Q-Learning process is significantly higher than the learning rate of the Semantic Representations Q-Learning process. The alpha distribution of the Choices Q-Learning process is more similar to that of the Semantic Representations Q-Learning process than of the Affective Representations Q-Learning process.

Table 1
The results of the mixed effects regressions in Experiment No. 1.

Dependent variable	Feeling - focused self -reports				Knowledge - focused self -reports			
Independent variable	Q -Values	Prediction Error	Q-Values	Prediction Error	Q -Values	Prediction Error	Q-Values	Prediction Error
Framework	Common accumulation		Parallel accumulation		Common accumulation		Parallel accumulation	
Q-Learning process	Choice Learner		Affective Representations Learner		Choice Learner		Semantic Representations Learner	
Intercept standardized coefficient	-0.02	0.002	-0.02	0.01	0.004	0.003	-0.01	<.001
Intercept Significance level	***	n.s	***	n.s	***	n.s	***	n.s
Independent variable standardized coefficient	0.38	0.26	0.43	0.22	0.72	0.19	0.78	0.18
Independent variable Significance level	***	***	**	***	***	***	***	***
AIC	7234	7669	7092	7765	6411	7800	6135	7832
BIC	7269	7705	7128	7800	6447	7835	6170	7867

2.3. Discussion

In experiment 1, we examined the effect of the expected (variability of rewards) and unexpected (reversal) uncertainty in the rewards on the affective and semantic valence learning processes and directly compared their temporal dynamics. As already apparent from the block-by-block dynamics of the two types of valence (Fig. 3.A), the affective valence ratings were updated faster than the semantic valence ratings and the choices. The different trend of correlations between current valence self–reports and the past US of the two types of valence gives the block–by–block descriptive result further support (Fig. 3.B). As reported above, the strength of the correlation between the current valence ratings and the short-term history of rewards declines in the case of the affective valence but not in the case of the semantic valence. This difference in correlations, which implies the higher weight given by the semantic valence to past events, converges with the slower adjustment of the semantic valence manifested in the block–by–block trends.

To formally analyze the differences between the two types of valence learning rates, we fitted three Q-Learning algorithms to the participant's trial-by-trial data: the *Choice Learner*, the *Affective Representations Learner*, and the *Semantic Representations Learner*. We then checked which of the two conceptual frameworks for learning best fit our data – the *Common Accumulation* framework or the *Parallel Accumulation* framework. We found that the *Parallel Accumulation* framework, which assumes a separate accumulation process for each type of valence, is a better fit for our data and proceeded under this framework.

Finally, we found that the estimated learning rate of the *Affective Representations Learner* process is significantly higher than the learning rate of the *Semantic Representations Learner* process and the *Choice Learner* process. The higher learning rate of the affective valence compared to the semantic valence (Fig. 3.C) implies that the affective valence representation adjustment process gives more weight to the short-term history of events, at the expense of its weight in the long–term history of events. In contrast, the semantic valence representation adjustment process gives more weight to the long-term history of events at the expense of the importance it gives to the short–term history of events. This difference in the adjustment dynamics of the two types of valence causes the faster affective valence adjustment than the adjustment of the semantic valence.

Notably, the lower learning rate of the semantic valence found in the first experiment may depend on this experiment's specific reward environment, i.e., the expected uncertainty caused by the variability of rewards connected to a specific CS. This variability causes a difference between the accumulated knowledge on the CS value based on different time windows (e.g., one trial back versus five trials back). This state results in a lack of accuracy of the CS predicted value and delayed stability in predictions. One way to mitigate these adverse effects of variability on the learning process is to consider the long-term history of events, i.e., use a low learning rate. Indeed, Preuschoff and Bossaerts (2007) and Diederen and Schultz (2015) show that in cases that involve expected variability, it is advantageous to adjust the prediction error based on the expected variability in rewards. The higher the variability the smaller the prediction error, and vice versa. This adjustment is computationally translated to a lower leaning rate the higher the variability of the rewards. The results of experiment 1 may suggest that the semantic valence representation attempted to address the expected (and unexpected) uncertainty in the environment by lowering the learning rate and smoothing out the trial-by-trial fluctuations in the rewards. In contrast, these fluctuations influenced the affective valence representation more strongly.

3. Experiment 2

In experiment 2, we aimed to check whether the difference between the learning rates of the affective and semantic valence representations stems from the expected variability in the rewards that were part of the design of experiment 1. To this end, we deliberately manipulated the **variability** of the reward schedule as a between participants' condition (see Supplementary Table 2). One group of participants, allocated to the "Variable rewards" condition, replicated experiment 1. During the acquisition phase (first 30 trials), choosing face "A" yielded an average profit of 10 points.[10] However, the actual points varied from 5 to 15, and selecting face "B" yielded an average of 0 points, but actual points varied from −5 to 5. The other group of participants, allocated to the "Fixed rewards" condition, experienced a highly predictive environment with no expected uncertainty, i.e., during the acquisition phase choosing face "A" yielded a fixed loss of 5 points and selecting face "B" yielded a fixed gain of 1 point. The contingencies in both groups, the Variable rewards condition and the Fixed reward condition, were reversed during the reversal phase (last 30 trials), as in the original experiment.

3.1. Method and material

Experiment 2 replicated and extended experiment 1 using the internet platform. Below, we will elaborate only on the differences relative to the first experiment.

3.1.1. Participants

For the second experiment, we recruited 109 participants from English-speaking countries (mainly the UK), using the Prolific platform [43 male, aged 34.3 ± 10.2 (mean ± SD); 66 female, aged 24.0 ± 11.2]. The study was approved by the University of Haifa's Faculty of Social Sciences Research Ethics Committee (Project ID Number: 318/18). All participants were screened for learning disabilities and attention deficits. Participants received GBP 8 for their participation. In addition, the participants received a bonus of up to GBP 3, depending on the points they received in 2 randomly selected trials and a fixed initial allocation of points. The average bonus was about 1.8 pounds. As elaborated in section 2.2.1 (Exclusion criteria), we excluded 29 participants from the study due to lack of learning (5 participants) or lack of sufficient variance in their self-report ratings (24 participants). The final sample size for the analysis was $n = 80$, which was randomly allocated to the two conditions, the Variable rewards condition and the Fixed rewards condition, i.e., $n = 40$ in each condition. The two final samples did not significantly differ in their demographic characteristics (i.e., mean age, percentage of females). Notably, we repeated the primary analysis presented below of comparing the learning rates of the two types of valence without exclusion of any participants and received similar results (See supplementary file section C for details).

3.2. Results

3.2.1. Cross-correlation analysis

Similar to the analysis performed in experiment 1 (section 2.2.3), we computed correlations between the current trial's valence self-report and the history of rewards (USs) for the two types of valence. The correlation calculations for each condition (i.e., Variable or Fixed rewards) in each time point (i.e., 0-lag, 1-lag, 2-lag, and 3-lag) were done on the individual participant's level, resulting in a distribution of 40 correlations. Each of these correlations was based on the correlation between 60 ratings and rewards related to a specific participant at a specific time. We then conducted a Fisher's Z transformation on the correlations.

As illustrated in Fig. 4.A, in the Variable rewards condition, the correlations between the feeling-focused self-reports and past rewards weakened as the disparity between the self-report and the time the reward was presented increased. A comparison between the two extremes (0-lag cross-correlation and 3-lag cross-correlation) showed a significant drop between the 0-lag cross-correlation ($M = 0.83$, $SD =$

[10] The participants were informed in advance that the task points would be converted to pounds at a ratio of 1 point equals 0.075 pounds.

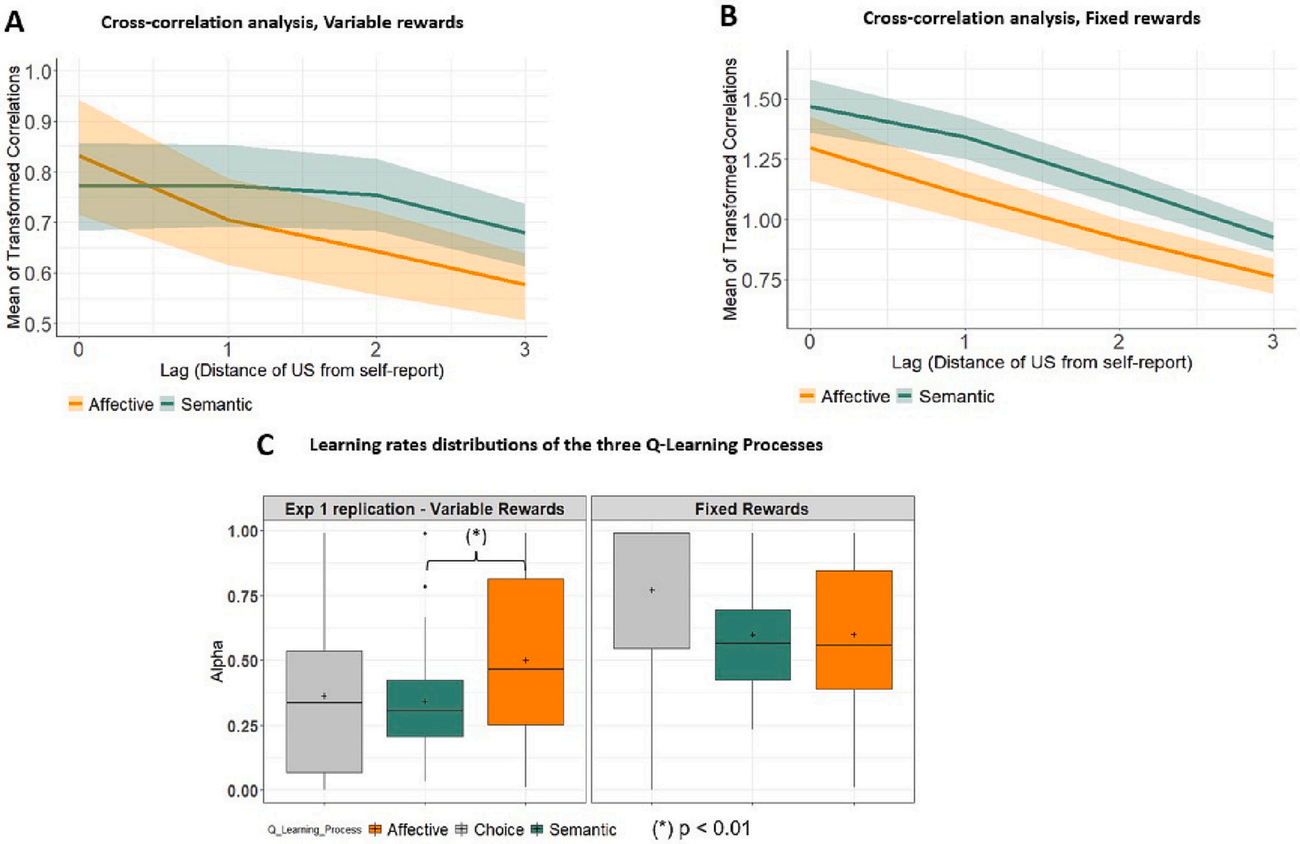

Fig. 4. Experiment 2 results. 4.A. Variable rewards condition (replication of experiment 1). The correlation between the affective valence self-report and the rewards (USs) weakens as the gap between the self-report and the time the US has been presented increases. The correlation between the semantic valence self-report and the US also weakens but to a lesser extent. 4.B Fixed rewards condition. The two types of self-report show a similar pattern of a weakening correlation between the self-reports and the US, as the gap between the self-report and the time the US was presented increased. 4.C Similar to the results of experiment 1, in the variable reward condition (replication) the alpha of the Affective Representations Q-Learning process is significantly higher than the alpha of the Semantic Representations Q-Learning process. In contrast, the alpha of the Affective Representations Q-Learning process in the Fixed rewards condition is not significantly different from the alpha of the Semantic Representations Q-Learning process.

0.38) and the 3-lag cross-correlation ($M = 0.58$, $SD = 0.21$) ($t(39) = 5.1$, $p < .0001$, $d = 0.80$). The correlation between the knowledge-focused self-reports and the rewards also decreased as the disparity between the self-reports and the time the rewards were presented increased, but to a lesser extent. A comparison between the 0-lag cross-correlation ($M = 0.77$, $SD = 0.28$) and 3-lag cross-correlation ($M = 0.68$, $SD = 0.20$) showed a significant drop ($t(39) = 2.39$, $p = .02$, $d = 0.38$).

In the Fixed rewards condition, the two types of valence showed a similar pattern. As illustrated in Fig. 4.B, the Pearson correlation between the feeling-focused self-reports and the rewards weakened as the gap between the self-report and the time in which the rewards were presented increased. A comparison between the 0-lag cross-correlation ($M = 1.30$, $SD = 0.47$) and 3-lag cross-correlation ($M = 0.76$, $SD = 0.24$) showed a significant drop ($t(39) = 8.24$, $p < .0001$, $d = 1.30$). The correlation between the knowledge-focused self-reports and the rewards also weakened as the disparity between the self-reports and the time in which the rewards were presented increased. A comparison between the 0-lag cross-correlation ($M = 1.47$, $SD = 0.38$) and 3-lag cross-correlation ($M = 0.92$, $SD = 0.21$) showed a significant drop ($t(39) = 10.30$, $p < .0001$, $d = 1.63$).

3.2.2. *Computational modeling*

3.2.2.1. Frameworks comparison – Variable rewards condition (replication). As detailed in the analysis of experiment 1 (see section 2.2.4.1), we started by fitting the three Q-Learning processes (i.e., the *Choice Learner*, the *Affective Representations Learner*, and the *Semantic Representations Learner*) to the trial-by-trial choices and ratings of the 40 participants. We obtained estimated parameters (e.g., learning rates) for each participant. We also recovered the trial-by-trial Q-Values and prediction errors for each learning process and used them for the frameworks comparison, as detailed below.

Table 2 below summarizes the results of eight mixed effects regression models with the affective (/semantic) valence self-reports ratings as the dependent variable for the Variable rewards condition. As in experiment 1, the best single variable to explain the feeling–focused self-reports is the Q-Values from the *Affective Representations Learner* process, which is our prediction to best fit under the *Parallel Accumulation* framework. Under the *Common Accumulation* framework, the best-explaining variable is the Q -Values of the *Choice Learner* process. Notably, the gap in AIC/BIC terms between the Q -Values from the *Affective Representations Learner* process and the Q -Values from the *Choice Learner* process is 5%. As in experiment 1, this difference can be attributed to the Affective Representations Learner being directly fitted to the feeling–focused self-reports. In contrast, the Choice Learner was fitted to the choices. For full details of the models, see supplementary tables D.9 to D.12.

The best single variable to explain the knowledge–focused self-reports is, again, like in experiment 1, our prediction for the best fit under the *Parallel Accumulation* framework, i.e., the Q-Values from the *Semantic Representations Learner* process. Under the *Common Accumulation* framework, the best-explaining variable is again the Q-Values from the *Choice Learner* process. The gap in AIC/BIC terms between the Q

Table 2
The results of the mixed effects regressions in Experiment No. 2.a – Variable rewards.

Dependent variable	Feeling - focused self -reports				Knowledge - focused self -reports			
Independent variable	Q -Values	Prediction Error	Q-Values	Prediction Error	Q -Values	Prediction Error	Q-Values	Prediction Error
Framework	Common accumulation		Parallel accumulation		Common accumulation		Parallel accumulation	
Q-Learning process	Choice Learner		Affective Representations Learner		Choice Learner		Semantic Representations Learner	
Intercept standardized coefficient	0.02	-0.008	0.004	-0.002	0.05	-0.004	0.01	-0.001
Intercept Significance level	***	n.s	***	n.s	***	n.s	***	n.s
Independent variable standardized coefficient	0.68	0.27	0.76	0.25	0.85	0.19	0.83	0.19
Independent variable Significance level	***	***	***	***	***	***	***	***
AIC	5797	6523	5524	6614	5433	6656	5073	6687
BIC	5832	6558	5559	6648	5467	6691	5107	6722

-Values from the *Semantic Representations Learner* and the Q -Values from the *Choice Learner* is 7%. This result gives some support to the *Parallel Accumulation* framework. For full details of the models, see supplementary tables D.13 to D.16.

In summary, like in experiment 1, the differences between the frameworks are minor and can be at least partly attributed to the direct fit of the self-reports to the learning processes in the *Parallel Accumulation* framework. However, it should be noted that again the *Common Accumulation* framework indicated that both types of self-reports are based on the same component of the *Choice Learner* process, i.e., its Q-Values, and therefore should be the same. As in experiment 1, this result contradicts our data that shows differences between the knowledge–focused and feeling–focused self–reports in both the block-to-block trends and the cross-correlation analysis. Therefore, we concluded that the *Parallel Accumulation* framework received more support from our data and chose it as the framework we adopted for the learning rates comparison detailed below.

3.2.2.2. Frameworks comparison – Fixed rewards condition. We repeated the frameworks comparison process above detailed, this time to the Fixed rewards condition data. The results are summarized in Table 3 below. As in experiment 1 and the Variable rewards condition of experiment 2, the best single variable to explain the feeling–focused self-reports is our prediction for the best fit under the *Parallel Accumulation* framework, i.e., the Q-Values from the *Affective Representations Learner* process. Under the *Common Accumulation* framework, the best-explaining variable is the Q -Values of the *Choice Learner* process. Notably, the gap in AIC/BIC terms between the Q -Values from the *Affective Representations Learner* process and the Q -Values from the *Choice Learner* process is 4%. As in experiment 1 and the Variable rewards condition of experiment 2, this difference can be attributed to the *Affective Representations* Learner being directly fitted to the feeling–focused self-reports. For full details of the models, see supplementary tables D.17 to D.20.

The best single variable to explain the knowledge–focused self-reports is, again, like in experiment 1 and the Variable rewards condition of experiment 2, our prediction for the best fit under the *Parallel Accumulation* framework, i.e., the Q-Values from the *Semantic*

Table 3
The results of the mixed effects regressions in Experiment No. 2.b – Fixed rewards.

Dependent variable	Feeling - focused self -reports				Knowledge - focused self -reports			
Independent variable	Q -Values	Prediction Error	Q-Values	Prediction Error	Q -Values	Prediction Error	Q-Values	Prediction Error
Framework	Common accumulation		Parallel accumulation		Common accumulation		Parallel accumulation	
Q-Learning process	Choice Learner		Affective Representations Learner		Choice Learner		Semantic Representations Learner	
Intercept standardized coefficient	-0.01	-0.001	-0.01	<.001	-0.01	-0.003	-0.01	<.001
Intercept Significance level	***	n.s	***	n.s	***	n.s	***	n.s
Independent variable standardized coefficient	0.76	0.19	0.80	0.20	0.84	0.16	0.88	0.17
Independent variable Significance level	***	***	***	***	***	***	***	***
AIC	4576	6687	4413	6690	3785	6698	3483	6708
BIC	4611	6722	4447	6725	3820	6732	3517	6743

Representations Learner process. Under the *Common Accumulation* framework, the best-explaining variable is again the Q-Values from the *Choice Learner* process. The gap in AIC/BIC terms between the Q -Values from the *Semantic Representations Learner* and the Q -Values from the *Choice Learner* is 9%. This result gives support to the *Parallel Accumulation* framework. For full details of the models, see supplementary tables D.21 to D.24.

In summary, in the Fixed rewards condition of experiment 2, as in experiment 1 and the Variable rewards condition of experiment 2, the *Parallel Accumulation* framework received more support. Again, the differences can be at least partly attributed to the direct fitting of the self-reports in the learning processes under the *Parallel Accumulation* framework. In addition, the *Common Accumulation* framework once again indicated that both types of self-reports are based on the same component of the *Choice Learner* process, i.e., its Q-Values, and therefore should be the same. Notably, this result does not contradict our data in the Fixed rewards case. Therefore, we conclude that both frameworks fit our data similarly. We decided to proceed under the *Parallel Accumulation* framework only to facilitate easy comparison to the Variable rewards condition. The learning rates comparison below supports the fact that the two allegedly separate learning processes of the two types of valence representations can be combined in the Fixed condition.

3.2.2.3. Comparing the learning rates of the semantic and affective valence. We moved to check whether the affective valence's learning rate differs from the semantic valence's learning rate under the *Parallel Accumulation* framework. To this end, we compared the estimated alpha parameters (i.e., the learning rate) derived from the *Affective Representations Learner* and the *Semantic Representations Learner* processes. In the Variable rewards condition (replication of experiment 1) we found that the estimated alpha from the *Affective Representations Learner* process ($M = 0.5$, $SD = 0.32$) is higher than the estimated alpha from the *Semantic Representations Learner* process ($M = 0.34$, $SD = 0.23$), and this difference is statistically significant ($t(39) = -2.92$, $p = .006$, $d = -0.46$) (also, see Fig. 4.C and supporting bayesian inference in supplementary fig. 5).[11] In contrast, in the Fixed rewards condition we found that the estimated alpha from the *Affective Representations Learner* process ($M = 0.6$, $SD = 0.28$) is not different from the estimated alpha from the *Semantic Representations Learner* processes ($M = 0.6$, $SD = 0.21$) ($t(39) = -0.02$, n.s, see supporting bayesian inference in supplementary fig. 6). We also conducted this analysis without excluding participants and found similar effects. See Supplementary section C for details.

Notably, the affective valence learning rate changed only marginally between the experiment's two conditions, from an estimated mean alpha of 0.5 in the Variable reward condition (and 0.48 in experiment 1) to an estimated mean alpha of 0.6 in the Fixed rewards condition. In contrast, the move from the Variable reward condition to the Fixed rewards condition strongly influenced the semantic valence. Its learning rate changed from an estimated mean alpha of 0.34 in the Variable reward condition (and 0.28 in experiment 1) to an estimated mean alpha of 0.6 in the Fixed rewards condition. This rise in the estimated mean alpha of the semantic valence in the Fixed reward condition closes the gap that exists between the two types of valence in the Variable reward condition.

3.3. Discussion

In experiment 2, we sought to directly examine the effect of the reward schedule's variability (i.e., expected uncertainty) on the existence of a difference between the temporal dynamics of the two types of valence. To this end, we deliberately manipulated the variability of the reinforcement schedule as a between participants' condition. One group of participants, allocated to the Variable reward condition, replicated experiment 1, i.e., experienced expected uncertainty; the variability of rewards was connected to each CS in each phase. The other group, allocated to the Fixed reward condition, experienced a fixed rewards schedule connected to each CS during each phase, i.e., did not experience expected uncertainty. Importantly, as in the design in experiment 1, the two groups experienced unexpected variability (i.e., reversal). In addition, as in experiment 1, our analysis was done under the *Parallel Accumulation* framework, which assumes three parallel Q-Learning processes, one for the choices and an additional two for each type of valence representation.

The results of the Variable rewards condition replicate experiment 1. In the less predictable environment of variability in the rewards schedule, the feeling-focused self-reports (i.e., affective valence) are updated faster than the knowledge-focused self-reports (i.e., semantic valence). This difference translates into a significantly higher affective valence learning rate than semantic valence learning rate. In contrast, in the highly predictable environment of the Fixed reward condition, the two types of valence have very similar temporal dynamics, manifested in their similar block–to–block patterns (Supplementary Fig. 3.B), correlations between current self-report and past rewards, and learning rates. Notably, as suggested by Preuschoff and Bossaerts (2007) and Diederen and Schultz (2015), the semantic valence representation mitigates the aversive effect of the unavoidable variability in the rewards (i.e., the expected uncertainty) on the learning process by scaling down the prediction error and by de facto lowering the learning rate. This mitigation causes dissociation between the temporal dynamics of the two types of valence in the face of the variability in rewards: The affective valence representation undergoes a faster but less stable adjustment, whereas the semantic valence representation undergoes a slower but smoother adjustment.

The difference between the variable and fixed rewards conditions is also reflected in the choice dynamics. Learning is easy and fast in the Fixed reward condition, a highly predictable environment. As a result, the participants learn to constantly choose the more profitable option relatively quickly, as displayed by the high learning rate. In contrast, learning is less trivial in the Variable rewards condition, resulting in the longer time it takes the participants to constantly choose the more profitable option.

4. General discussion

The current study investigated the temporal dynamics of affective and semantic representations of valence in a reversal-learning task. Experiment 1 implements a schedule of variable rewards; the participants are exposed to both expected uncertainty (i.e., uncertainty in the rewards associated with each CS in each phase) and unexpected uncertainty, i.e., the reversal. We found that in this environment, the learning rate of the choices and the knowledge-focused self-reports is lower than the learning rate of the feeling-focused self-reports. In other words, we found that in an environment that contains both expected and unexpected uncertainty regarding rewards, the semantic valence representations of the CSs are updated at a slower pace than the affective valence representations of the CSs.

In the second experiment, we implemented two conditions. The first is the Variable rewards condition in which, as in experiment 1, the participants experienced both expected and unexpected uncertainty. The second is the Fixed rewards condition, in which the participants experienced *only* unexpected uncertainty, i.e., a reversal. The results of the Variable rewards condition replicate experiment 1. In the presence of variability in the rewards schedule associated with each CS in each phase and a reversal, the learning rate of the choices and the knowledge-focused self-reports are lower than that of feeling–focused self–reports. In other words, we again found that in an environment that contains both expected and unexpected uncertainty regarding the rewards, the semantic valence representations of the CSs are updated at a slower pace

[11] For additional information on the results of the Q-Learning processes see supplementary table 3.

than the affective valence representations of the CSs. In contrast, in the Fixed rewards condition (i.e., only unexpected uncertainty, no expected uncertainty), there is no difference between the update pace, i.e., the learning rates of the two types of valence representations, as they are both relatively high.

Notably, the difference between the results of the Variable and Fixed rewards conditions stems from the different effects of expected uncertainty on the learning rates of the two types of valence. The semantic valence responds to expected uncertainty by significantly decreasing its learning rate. On the other hand, the affective valence is only marginally lower with expected uncertainty. Combining the two different trends generates a discrepancy between the learning rates of the two types of valence in the Variable rewards (i.e., expected and unexpected uncertainty) condition that does not exist in the Fixed rewards condition (i.e., only unexpected uncertainty). We conclude that the semantic valence representations for the CSs are updated slower than the affective valence representations of the CSs, i.e., they have a lower learning rate *only* in an environment that contains expected uncertainty.

A valid question is why expected uncertainty might lower the learning rate, as we primarily found in the case of the semantic valence representation. To answer this question, we should first recall that expected uncertainty is the uncertainty in rewards attributable to the inherent variability of the given phenomenon. Therefore, if there is expected uncertainty, the current reward is not exactly identical to the reward in the last trial and the reward two trials ago. In other words, if there is expected uncertainty, the history of rewards is informative because it is not exactly like the current reward. On the other hand, no expected uncertainty means that the history of rewards is far less informative because the current reward is the same as the last trial's reward and the same as the reward two trials ago. Let us now recall that the learning rate controls the weight given to the long-term reward history versus the current reward in determining the CS's valence. The higher the learning rate, the less weight given to the long-term reward history and the more weight given to the current reward. As explained above, if there is no expected uncertainty, the long-term reward history is far less informative. Therefore, both types of valence representations can adopt a relatively high learning rate, primarily relying on the current reward and not on the long-term reward history. On the other hand, if there is expected uncertainty, the long-term reward history is informative. In this case, adopting a high learning rate means that the current reward, with its inherited fluctuations, will significantly influence the CS's value, causing the adaptation process to be faster but more volatile. In contrast, adopting a low learning rate means that the long-term reward history and not the current reward will primarily influence the CS's value, causing the adaptation process to be slower but more stable (/smoother). We found that in the face of expected uncertainty, the semantic valence representation (and the choices) preferred adopting a lower learning rate than the affective valence representation, causing a slower but smoother adaptation process of the semantic valence representation than the affective valence representation.

Our results support the dissociation between the semantic and affective representations of valence. This finding has implications for several fields of knowledge. First, for basic emotion research. We found additional empirical evidence for a dissociation between the two valence systems that are usually connected and treated in the literature as one construct. The de facto treatment of affective and semantic representations of valence as a monolithic structure stems from the challenge of dissociating them. It creates unfavorable consequences for basic emotion theory and empirical work. For example, it can lead to misuse of self–reports and tasks, e.g., using "affective" self-reports that reflect semantic evaluation; using "affective" tasks that reflect semantic components (see Itkes & Kron, 2019, for a review). The dissociation between the two types of valence in the face of expected uncertainty, i.e., the variability of rewards, contributes to our understanding of the difference between affective and semantic valence representations. The semantic valence representation focuses more on the long-term history of events, whereas the affective valence representation focuses more on current events.

A second implication is contributing to learning theories by considering semantic and affective learning simultaneously. The lion's share of the existing literature on associative learning deals with either value-based learning, i.e., learning about the semantic valence of a stimulus (e.g., Boldt, Blundell, & De Martino, 2019; Daw, Gershman, Seymour, Dayan, & Dolan, 2011; Hertz et al., 2018) or learning about the affective valence of a stimulus, i.e., fear conditioning (e.g., Atlas et al., 2016; Mertens & De Houwer, 2016; Schiller et al., 2008). We extended previous work by considering semantic valence, affective valence, and choices simultaneously. To achieve this goal, we first checked which conceptual learning framework best fit our data by formally comparing two frameworks. According to the first framework, there is a single learning process with a single learning rate. The affective and semantic valence representations are based on the read-out of latent variables of this learning process. According to the second framework, there are three parallel accumulation processes, one for the choices and one for each type of valence representation, i.e., the same schedule of rewards (US) is used to create two different representations of the CS' valence, one of the semantic valence of the CS and one of the affective valence of the CS.

The *Parallel Accumulation* framework demonstrated a minor to moderate advantage in both experiments over the *Common Accumulation* framework. This advantage can be at least partly attributed to the direct fit of each type of self-report to a dedicated learning process, whereas the common accumulation model was fit to the choices. However, we predicted that the *Common Accumulation* framework could capture differences between semantic and affective ratings by linking the ratings to different variables – affective to prediction error and semantic to Q-values or vice-versa. Our results indicated that the prediction errors did not contribute much to the ratings made by participants, therefore rejecting a fundamental assumption in the common accumulator framework. As it stands, this framework indicated that both types of self-reports are based on the same component of the *Choice Learner* process, i.e., its Q-Values, and therefore should be the same. This result contradicts our data in experiment 1 and the variable rewards condition of experiment 2, which shows differences between the knowledge–focused and feeling–focused self–reports in both the block-to-block trends and the cross-correlation analysis. Therefore, we concluded that the *Parallel Accumulation* framework provided a more meaningful description of our results, even though it is more elaborate than the common accumulator farmework. We then directly compared the two types of valence learning processes under this framework, using the same Q-Learning algorithm for the two types of valence. It should be noted that future works using different experimental designs and reward patterns may be able to link ratings with the prediction error variable better.

Notably, our results support Preuschoff and Bossaerts (2007) and Diederen and Schultz (2015). They claim that in cases that involve variability in rewards, it is advantageous to adjust the prediction error based on the expected variability in rewards. The higher the variability, the smaller the prediction error and vice versa. This adjustment is computationally translated to a lower learning rate the higher the variability of the rewards. Indeed, our results show that in the face of variability in rewards, the semantic valence representation (and the choices) preferred adopting a low learning rate, causing a slower but smoother adaptation process.

Finally, our results have a potential contribution to research on affect and emotion in decision-making (for reviews, see Lerner, Li, Valdesolo, & Kassam, 2015, and Phelps, Lempert, & Sokol-Hessner, 2014). It is interesting to look at the connection between the valence representations' temporal dynamics and the temporal dynamics of choices, i.e., the decisions made by the participants. Our results are limited in providing a direct comparison between the learning rates estimated from the choices Q-Learning algorithm (i.e., the *Choice Learner)* and the learning rates estimated from the affective and semantic valence representations Q-

Learning algorithms (i.e., the *Affective Representations Learner* and the *Semantic Representations Learner*) due to the different error terms in these Q-Learning algorithms. Nevertheless, the descriptive block-by-block trends (see Fig. 3.A) and the learning rates (see Fig. 3.C and Fig. 4.C) imply that the choices' dynamics might be more similar to the semantic than to the affective valence dynamics. Therefore, we can speculate that the participants possibly relied more on semantic valence representation than on affective valence representation when making their decisions. Future research may wish to use the paradigm of associative learning to investigate the effect of different types of valence representations on decision-making.

Our primary finding, i.e., the faster adaptation of the emotional valence representation in volatile, complex environments, also resonates with the Somatic Marker Hypothesis (SMH) (Damasio, Tranel, & Damasio, 1991; Damasio, 1996; for critical review see Dunn, Dalgleish, & Lawrence, 2006). The hypothesis claims that emotion-based biasing signals affect decision-making in complex situations. These signals indicate our emotional reaction to the response option and are integrated into higher brain regions, particularly the ventromedial prefrontal cortex (VMPFC) (Bechara, Damasio, Damasio, & Anderson, 1994). In particular connection to the current study, Bechara and his colleagues suggest that explicit reasoning in cases that involve uncertainty is preceded by rapidly formed emotional-based nonconscious biasing that supports declarative knowledge and advantageous behavior (Bechara et al., 1997).

On this point, it is essential to distinguish between the basic assumptions adopted in this study and those adopted by the SMH and similar theories (e.g., Loewenstein, Hsee, Weber, & Welch, 2001; Zajonc, 1980). These theories consider emotions' output for decision-making as a heuristic, unconscious process that gives a quick but less accurate evaluation of the alternatives (i.e., affect heuristics view). We checked whether our data support such a fundamental difference between the affective valence representation and the semantic valence representation, i.e., is the affective valence representation based on prediction errors and the semantic valence representation based on accumulated knowledge (Q-Values). Despite the attractiveness of a parsimonious single learning process with different read-outs, our data supported the *parallel Accumulation* framework, which assumes three parallel accumulation processes, one for the choices and one for each type of valence representation. Under this framework, feelings, the conscious part of emotions, are not fundamentally different from semantic knowledge as an output for decision-making. Therefore, we modeled the two types of valence representations, the semantic and the affective, using the same formulation. Furthermore, note that in our design, affective valence is measured using feeling self-reports. The participants had to be aware of their feelings to report them, and consequently, the emotional contribution to the decision-making process had to be conscious. To conclude, our methodology and results support the view of the affective and semantic valence representations as two unique and interactive contributions to the decision-making process, similar to the conclusions of Quartz (2009) and Heffner, Son, and Feldmanhall's (2021).

One potential limitation of our studies is using neutral human faces as the CSs. Human subjects have a congenital tendency to develop emotions toward human faces (Grossmann, 2015). Therefore, the results of the experiments cannot be automatically generalized to other stimuli. Future research should use generic objects such as colored blocks as stimuli. More importantly, our dissociation results between the two types of valence representation in the case of variability of rewards stem from an investigation of only two conditions. In one condition, the rewards were fixed; in the other, the rewards were variable. Future research should seek to extend the analysis and uncover a more general functional connection between the variability of rewards and the difference between the learning rates of the two valence representations by systematically controlling the learning task's level of variability and examining the resulting difference between the two valence representations. Finally, our current findings supported a parallel accumulators framework, mainly because the prediction error variable seems not to play a significant role in predicting participants' ratings, resulting in two independent processes of information accumulation. Future works may use other experimental designs or neuroimaging techniques to track these separate processes and their neural basis and establish their independence.

To conclude, the current study provides direct evidence for the existence of two different learning processes of a stimulus' valence in an unstable rewards environment. In the face of variability, the semantic knowledge about the stimulus valence, i.e., the semantic valence representation, smooths the fluctuations in the rewards by updating slower than the valence of the affective response to the same stimulus, i.e., the affective valence representation. This finding supports past evidence for the existence of two different representations of a stimulus' valence. It also strengthens the need to distinguish between their unique effect on behavior in future research.

Funding

UH was supported by the National Institute of Psychobiology in Israel (211-19-20) and the Israel Science Foundation (1532/20). AK and OH were supported by the Israel Science Foundation (1134/18).

CRediT authorship contribution statement

Orit Heimer: Conceptualization, Software, Formal analysis, Writing – original draft, Writing – review & editing, Visualization. **Assaf Kron:** Conceptualization, Methodology, Writing – original draft, Writing – review & editing, Supervision, Funding acquisition. **Uri Hertz:** Conceptualization, Methodology, Writing – original draft, Writing – review & editing, Supervision, Funding acquisition.

Declaration of Competing Interest

The authors declare no conflict of interest concerning the publication or the authorship of this article.

Data availability

All data and analysis scripts are available at: https://osf.io/zt7jw/

Acknowledgments

We thank Idit PhD Fellowship for its support of this project.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.cognition.2023.105423.

References

Akaike, H. (1974). A new look at the statistical model identification. *IEEE Transactions on Automatic Control, 19*(6), 716–723.

Atlas, L. Y. (2019). How instructions shape aversive learning: Higher order knowledge, reversal learning, and the role of the amygdala. *Current Opinion in Behavioral Sciences, 26*, 121–129. https://doi.org/10.1016/j.cobeha.2018.12.008

Atlas, L. Y., Doll, B. B., Li, J., Daw, N. D., & Phelps, E. A. (2016). Instructed knowledge shapes feedback- driven aversive learning in striatum and orbitofrontal cortex, but not the amygdala. *ELife, 5*(MAY2016), 1–26. https://doi.org/10.7554/eLife.15192

Barrett, L. F. (2006a). Solving the emotion paradox: Categorization and the experience of emotion. *Personality and Social Psychology Review, 10*(1), 20–46. https://doi.org/10.1207/s15327957pspr1001_2

Barrett, L. F. (2006b). Valence is a basic building block of emotional life. *Journal of Research in Personality, 40*(1), 35–55. https://doi.org/10.1016/j.jrp.2005.08.006

Bechara, A., Damasio, A. R., Damasio, H., & Anderson, S. W. (1994). Insensitivity to future consequences following damage to human prefrontal cortex. *Cognition, 50* (1–3), 7–15. https://doi.org/10.1016/0010-0277(94)90018-3

Bechara, A., Damasio, H., Tranel, D., Damasio, A. R., Bushnell, M. C., Matthews, P. M., & Rawlins, J. N. P. (1997). Deciding advantageously before knowing the advantageous strategy. *Science, 275*(5304), 1293–1295. https://doi.org/10.1126/science.275.5304.1293

Boddez, Y., Baeyens, F., Luyten, L., Vansteenwegen, D., Hermans, D., & Beckers, T. (2013). Rating data are underrated: Validity of US expectancy in human fear conditioning. *Journal of Behavior Therapy and Experimental Psychiatry, 44*(2), 201–206. https://doi.org/10.1016/j.jbtep.2012.08.003

Boldt, A., Blundell, C., & De Martino, B. (2019). Confidence modulates exploration and exploitation in value-based learning. *Neuroscience of Consciousness, 2019*(1), 1–12. https://doi.org/10.1093/nc/niz004

Cardinal, R. N., Parkinson, J. A., Hall, J., & Everitt, B. J. (2002). Emotion and motivation: The role of the amygdala, ventral striatum, and prefrontal cortex. *Neuroscience and Biobehavioral Reviews, 26*(3), 321–352. https://doi.org/10.1016/S0149-7634(02)00007-6

Carroll, J. M., Yik, M. S. M., Russell, J. A., & Barrett, L. F. (1999). On the psychometric principles of affect. *Review of General Psychology, 3*(1), 14–22. https://doi.org/10.1037/1089-2680.3.1.14

Codispoti, M., Ferrari, V., & Bradley, M. M. (2006). Repetitive picture processing: Autonomic and cortical correlates. *Brain Research, 1068*(1), 213–220. https://doi.org/10.1016/j.brainres.2005.11.009

Costa, V. D., Bradley, M. M., & Lang, P. J. (2015). From threat to safety: Instructed reversal of defensive reactions. *Psychophysiology, 52*(3), 325–332. https://doi.org/10.1111/psyp.12359

Dael, N., Mortillaro, M., & Scherer, K. R. (2012). Emotion expression in body action and posture. *Emotion, 12*(5), 1085–1101. https://doi.org/10.1037/a0025737

Damasio, A. R. (1996). The somatic marker hypothesis and the possible functions of the prefrontal cortex. *Philosophical Transactions of the Royal Society of London, Series B, 351*(1346), 1413–1420.

Damasio, A. R., Tranel, D., Damasio, H., Eisenberg, H. S., & Benton, H. M. (1991). Somatic markers and the guidance of behaviour: Theory and preliminary testing. In A. L. Levin (Ed.), *Frontal lobe function and dysfunction. Oxford* (pp. 217–229). Damasio,). Oxford University Press.

Daw, N. D., & Doya, K. (2006). The computational neurobiology of learning and reward. *Current Opinion in Neurobiology, 16*(2), 199–204. https://doi.org/10.1016/j.conb.2006.03.006

Daw, N. D., Gershman, S. J., Seymour, B., Dayan, P., & Dolan, R. J. (2011). Model-based influences on humans' choices and striatal prediction errors. *Neuron, 69*(6), 1204–1215. https://doi.org/10.1016/j.neuron.2011.02.027

Daw, N. D., & O'Doherty, J. P. (2014). Multiple systems for value learning. In *Neuroeconomics: Decision making and the brain* (Second edition, pp. 393–410). https://doi.org/10.1016/B978-0-12-416008-8.00021-8

De Houwer, J., & Eelen, P. (1998). An affective variant of the Simon paradigm. *Cognition & Emotion, 2*(1), 45–61. http://apps.webofknowledge.com.libproxy.ucl.ac.uk/full_record.do?product=UA&search_mode=GeneralSearch&qid=45&SID=N2fE9DYXQ6TZqDrTvw9&page=1&doc=4.

Delgado, M. R., Labouliere, C. D., & Phelps, E. A. (2006). Fear of losing money? Aversive conditioning with secondary reinforcers. *Social Cognitive and Affective Neuroscience, 1*(3), 250–259. https://doi.org/10.1093/scan/nsl025

Deutsch, R., & Strack, F. (2004). Reflective and impulsive determinants of social behavior. *Personality and Social Psychology Review, 8*(3), 220–247 (internal-pdf:/strack & roland (2004) RIM.pdf).

Diederen, K. M. J., & Schultz, W. (2015). Scaling prediction errors to reward variability benefits error-driven learning in humans. *Journal of Neurophysiology, 114*(3), 1628–1640. https://doi.org/10.1152/jn.00483.2015

Dunn, B. D., Dalgleish, T., & Lawrence, A. D. (2006). The somatic marker hypothesis: A critical evaluation. *Neuroscience and Biobehavioral Reviews, 30*(2), 239–271. https://doi.org/10.1016/j.neubiorev.2005.07.001

Duscherer, K., Holender, D., & Molenaar, E. (2008). Revisiting the affective Simon effect. *Cognition and Emotion, 22*(2), 193–217. https://doi.org/10.1080/02699930701339228

Erdfelder, E., FAul, F., Buchner, A., & Lang, A. G. (2009). Statistical power analyses using G*power 3.1: Tests for correlation and regression analyses. *Behavior Research Methods, 41*(4), 1149–1160. https://doi.org/10.3758/BRM.41.4.1149

Garrido, M. V., & Prada, M. (2017). KDEF-PT: Valence, emotional intensity, familiarity and attractiveness ratings of angry, neutral, and happy faces. *Frontiers in Psychology, 8*, 2181. https://doi.org/10.3389/fpsyg.2017.02181

Gazzaniga, M. S., Ivry, R. B., & Mangun, G. R. (2014). *Cognitive neuroscience : The biology of the mind (fourth Edi)*. W.W.Norton.

Gelman, A., & Hill, J. (2006). *Data analysis using regression and multilevel/hierarchical models*. Cambridge University Press.

Givon, E., Itzhak-Raz, A., Karmon-Presser, A., Danieli, G., & Meiran, N. (2019). How does the emotional experience evolve? Feeling generation as evidence accumulation. *Emotion, 20*(2), 271–285. https://doi.org/10.1037/emo0000537

Gläscher, J., & Büchel, C. (2005). Formal learning theory dissociates brain regions with different temporal integration. *Neuron, 47*(2), 295–306. https://doi.org/10.1016/j.neuron.2005.06.008

Goeleven, E., De Raedt, R., Leyman, L., & Verschuere, B. (2008). The Karolinska directed emotional faces : A validation study. *Cognition & Emotion, 22*(6). https://doi.org/10.1080/02699930701626582

Green, L., & Myerson, J. (2004). A discounting framework for choice with delayed and probabilistic rewards. *Psychological Bulletin, 130*(5), 769.

Grossmann, T. (2015). The development of social brain functions in infancy. *Psychological Bulletin, 141*(6), 1266–1287. https://doi.org/10.1037/bul0000002

Haj-Ali, H., Anderson, A. K., & Kron, A. (2020). Comparing three models of arousal in the human brain. *Social Cognitive and Affective Neuroscience, 15*(1), 1–11. https://doi.org/10.1093/scan/nsaa012

Hamzani, O., Mazar, T., Itkes, O., Petranker, R., & Kron, A. (2019). Semantic and affective representations of valence: Prediction of autonomic and facial responses from feelings-focused and knowledge-focused self-reports. *Emotion, 20*(3), 486–500. https://doi.org/10.1037/emo0000567

Heffner, J., Son, J., & Feldmanhall, O. (2021). Emotion prediction errors guide socially adaptive behaviour. *Nature Human Behaviour, 5*, 1391–1401. https://doi.org/10.1038/s41562-021-01213-6

Hertz, U., Bahrami, B., & Keramati, M. (2018). Stochastic satisficing account of confidence in uncertain value-based decisions. *PLoS One, 13*(4), 1–23.

Hodes, R. L., Cook, E. W., III, & Lang, P. J. (1985). Individual differences in autonomic response: Conditioned association or conditioned fear? *Psychophysiology, 22*(5), 545–560.

Houtveen, J. H., Rietveld, S., Schoutrop, M., Spiering, M., & Brosschot, J. F. (2001). A repressive coping style and affective, facial and physiological responses to looking at emotional pictures. *International Journal of Psychophysiology, 42*(3), 265–277. https://doi.org/10.1016/S0167-8760(01)00150-7

Itkes, O., Kimchi, R., Haj-Ali, H., Shapiro, A., & Kron, A. (2017). Dissociating affective and semantic valence. *Journal of Experimental Psychology: General, 146*(7), 924–942. https://doi.org/10.1037/xge0000291

Itkes, O., & Kron, A. (2019). Affective and semantic representations of valence: A conceptual framework. *Emotion Review, 11*(4), 283–293. https://doi.org/10.1177/1754073919868759

Izquierdo, A., Brigman, J. L., Radke, A. K., Rudebeck, P. H., & Holmes, A. (2017). The neural basis of reversal learning: An updated perspective. *Neuroscience, 345*, 12–26. https://doi.org/10.1016/j.neuroscience.2016.03.021

Kringelbach, M. L., O'Doherty, J., Rolls, E. T., & Andrews, C. (2003). Activation of the human orbitofrontal cortex to a liquid food stimulus is correlated with its subjective pleasantness. *Cerebral Cortex, 13*(10), 1064–1071.

Kron, A., Goldstein, A., Lee, D. H. J., Gardhouse, K., & Anderson, A. K. (2013). How are you feeling? Revisiting the quantification of emotional qualia. *Psychological Science, 24*(8), 1503–1511. https://doi.org/10.1177/0956797613475456

Kron, A., & Weksler, A. (2022). The feelings of goals hypothesis: Emotional feelings are non-conceptual, Non-Motoric Representations of Goals. *Emotion Review*. https://doi.org/10.1177/17540739221104456

Kross, E., & Ayduk, O. (2011). Making meaning out of negative experiences by self-distancing. *Current Directions in Psychological Science, 20*(3), 187–191. https://doi.org/10.1177/0963721411408883

Lang, P. J., Bradley, M. M., & Cuthbert, B. N. (1997). International affective picture system (IAPS): Technical manual and affective ratings. *NIMH Center for the Study of Emotion and Attention, 1*(39–58), 3.

Larsen, J. T., Norris, C. J., & Cacioppo, J. T. (2003). Effects of positive and negative affect on electromyographic activity over zygomaticus major and corrugator supercilii. *Psychophysiology, 40*(5), 776–785. https://doi.org/10.1111/1469-8986.00078

Lazarus, R. S., & Smith, C. A. (1988). Knowledge and appraisal in the cognition-emotion relationship. *Cognition and Emotion, 2*(4), 281–300. https://doi.org/10.1080/02699938808412701

Lempert, K., & Phelps, A. (2014). Neuroeconomics of emotion and decision making. In P. W. Glimcher, & E. Fehr (Eds.), *Neuroeconomics: Decision making and the brain* (2nd ed., pp. 219–236). Amsterdam: Academic Press.

Lerner, J. S., Li, Y., Valdesolo, P., & Kassam, K. S. (2015). Emotion and decision making. *Annual Review of Psychology, 66*, 799–823. https://doi.org/10.1146/annurev-psych-010213-115043

Lewis, M. D. (2005). Bridging emotion theory and neurobiology through dynamic systems modeling. *The Behavioral and Brain Sciences, 28*(2), 169–194. discussion 194-245. http://www.ncbi.nlm.nih.gov/pubmed/16201458

Loewenstein, G. F., Hsee, C. K., Weber, E. U., & Welch, N. (2001). Risk as feelings. *Psychological Bulletin, 127*(2), 267–286. https://doi.org/10.1037/0033-2909.127.2.267

Lundqvist, D., Flykt, A., & Öhman, A. (1998). The Karolinska directed emotional faces – KDEF, CD ROM from Department of Clinical Neuroscience, Psychology Section. *Karolinska Institutet*, 91–630.

Mertens, G., & De Houwer, J. (2016). Potentiation of the startle reflex is in line with contingency reversal instructions rather than the conditioning history. *Biological Psychology, 113*, 91–99. https://doi.org/10.1016/j.biopsycho.2015.11.014

Metcalfe, J., & Mischel, W. (1999). A hot/cool-system analysis of delay of gratification: Dynamics of willpower. *Psychological Review, 106*(1), 3–19. https://doi.org/10.1037/0033-295X.106.1.3

Mitchell, C. J., De Houwer, J., & Lovibond, P. F. (2009). The propositional nature of human associative learning. *Behavioral and Brain Sciences, 32*(2), 183–246. https://doi.org/10.1017/S0140525X09000855

Niv, Y. (2009). Reinforcement learning in the brain. *Journal of Mathematical Psychology, 53*(3), 139–154. https://doi.org/10.1016/j.jmp.2008.12.005

Osgood, C. E. (1952). The nature and measurement of meaning. *Psychological Bulletin, 49*(3), 197.

Osgood, C. E., Suci, G. J., & Tannenbaum, P. H. (1957). *The measurement of meaning (issue 47)*. University of Illinois press.

Peters, J., & Büchel, C. (2010). Neural representations of subjective reward value. *Behavioural Brain Research, 213*(2), 135–141. https://doi.org/10.1016/j.bbr.2010.04.031

Phelps, E. A., Lempert, K. M., & Sokol-Hessner, P. (2014). Emotion and decision making: Multiple modulatory neural circuits. *Annual Review of Neuroscience, 37*(1), 263–287. https://doi.org/10.1146/annurev-neuro-071013-014119

Preuschoff, K., & Bossaerts, P. (2007). Adding prediction risk to the theory of reward learning. *Annals of the New York Academy of Sciences, 1104*, 135–146. https://doi.org/10.1196/annals.1390.005

Quartz, S. R. (2009). Reason, emotion and decision-making: Risk and reward computation with feeling. *Trends in Cognitive Sciences, 13*(5), 209–215. https://doi.org/10.1016/j.tics.2009.02.003

Quilty-Dunn, J. (2016). Iconicity and the format of perception. *Journal of Consciousness Studies, 23*(3–4), 255–263.

Rangel, A., Camerer, C., & Montague, P. R. (2008). A framework for studying the neurobiology of value-based decision making. *Nature Reviews Neuroscience, 9*(7), 545–556. https://doi.org/10.1038/nrn2357

Rescorla, R. A., & Wagner, A. R. (1972). A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement. In A. H. Black, & W. F. Prokasy (Eds.), *Classical conditioning II: Current research and theory* (pp. 64–99). Appleton-Century-Crofts.

Robinson, M. D., & Clore, G. L. (2002a). Belief and feeling: Evidence for an accessibility model of emotional self-report. *Psychological Bulletin, 128*(6), 934–960. https://doi.org/10.1037/0033-2909.128.6.934

Robinson, M. D., & Clore, G. L. (2002b). Episodic and semantic knowledge in emotional self-report: Evidence for two judgment processes. *Journal of Personality and Social Psychology, 83*(1), 198–215. https://doi.org/10.1037/0022-3514.83.1.198

Roseman, I. J., & Smith, C. A. (2001). Appraisal Theory - Overview,Assumptions, Varieties, Controversies. In K. R. Scherer, A. Schorr, & T. Johnstone (Eds.), *Appraisal processes in emotion : Theory, methods, research* (pp. 3–19). Oxford: Oxford University Press.

Ruff, C. C., & Fehr, E. (2014). The neurobiology of rewards and values in social decision making. *Nature Reviews Neuroscience, 15*(8), 549–562. https://doi.org/10.1038/nrn3776

Russell, J. A. (1980). A circumplex model of affect. *Journal of Personality and Social Psychology, 39*(6), 1161–1178.

Russell, J. A. (1983). Pancultural aspects of the human conceptual organization of emotions. *Journal of Personality and Social Psychology, 45*(6), 1281–1288. https://doi.org/10.1037/0022-3514.45.6.1281

Russell, J. A. (2003). Core affect and the psychological construction of emotion. *Psychological Review, 110*(1), 145–172. https://doi.org/10.1037/0033-295X.110.1.145

Russell, J. A. (2005). Emotion in human consciousness is built on core affect. *Journal of Consciousness Studies, 12*(8–10), 26–42.

Russell, J. A., & Barrett, L. F. (1999). Core affect, prototypical emotional episodes, and other things called emotion: Dissecting the elephant. *Journal of Personality and Social Psychology, 76*(5), 805.

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 of the United States of America, 111*(33), 12252–12257. https://doi.org/10.1073/pnas.1407535111

Sander, D., Grandjean, D., Kaiser, S., Wehrle, T., & Scherer, K. R. (2007). Interaction effects of perceived gaze direction and dynamic facial expression: Evidence for appraisal theories of emotion. *European Journal of Cognitive Psychology, 19*(3), 470–480. https://doi.org/10.1080/09541440600757426

Schacter, D. L., Wagner, A. D., & Buckner, R. L. (2000). Memory systems of 1999. In *The Oxford handbook of memory* (pp. 627–643). Oxford University Press.

Schaefer, A., Collette, F., Philippot, P., Van Der Linden, M., Laureys, S., Delfiore, G., Degueldre, C., Maquet, P., Luxen, A., & Salmon, E. (2003). Neural correlates of "hot" and "cold" emotional processing: A multilevel approach to the functional anatomy of emotion. *NeuroImage, 18*(4), 938–949. https://doi.org/10.1016/S1053-8119(03)00009-0

Schiller, D., & Delgado, M. R. (2010). Overlapping neural systems mediating extinction, reversal and regulation of fear. *Trends in Cognitive Sciences, 14*(6), 268–276. https://doi.org/10.1016/j.tics.2010.04.002

Schiller, D., Levy, I., Niv, Y., LeDoux, J. E., & Phelps, E. A. (2008). From fear to safety and back: Reversal of fear in the human brain. *The Journal of Neuroscience, 28*(45), 11517–11525. https://doi.org/10.1523/JNEUROSCI.2265-08.2008

Schwarz, G. (1978). Estimating the dimension of a model. *The Annals of Statistics*, 461–464.

Smith, M. A., Ghazizadeh, A., & Shadmehr, R. (2006). Interacting adaptive processes with different timescales underlie short-term motor learning. *PLoS Biology, 4*(6), Article e179. https://doi.org/10.1371/journal.pbio.0040179

Soltani, A., & Izquierdo, A. (2019). Adaptive learning under expected and unexpected uncertainty. *Nature Reviews Neuroscience, 20*(10), 635–644. https://doi.org/10.1038/s41583-019-0180-y

Soltani, A., Murray, J. D., Seo, H., & Lee, D. (2021). Timescales of cognition in the brain. *Current Opinion in Behavioral Sciences, 41*, 30–37. https://doi.org/10.1016/j.cobeha.2021.03.003

Sutton, R. S., & Barto, A. G. (2018). *Reinforcement learning: An introduction* (2nd ed.). The MIT Press.

Tulving, E. (1983). *Elements of episodic memory*. Oxford University Press.

Tulving, E. (1993). What is episodic memory? *Current Directions in Psychological Science, 2*(3), 67–70. https://doi.org/10.1111/1467-8721.ep10770899

Wang, L., Li, X., Pi, Z., Xiang, S., Yao, X., & Qi, S. (2021). Spatiotemporal dynamics of affective and semantic valence among women. *Frontiers in Human Neuroscience, 15*, Article 602192. https://doi.org/10.3389/fnhum.2021.602192

Watkins, C. J. C. H. (1989). *Learning with delayed rewards* (Unpublished Ph.D.Thesis). Cambridge, UK: Cambridge University.

Wheeler, M. A., Stuss, D. T., & Tulving, E. (1997). Toward a theory of episodic memory: The frontal lobes and autonoetic consciousness. *Psychological Bulletin, 121*(3), 331–354. https://doi.org/10.1037/0033-2909.121.3.331

Zajonc, R. B. (1980). Feeling and thinking: Preferences need no inferences. *American Psychologist, 35*(2), 151–175. https://doi.org/10.1037/0003-066X.35.2.151

Zajonc, R. B. (1984). On the primacy of affect. *American Psychologist, 39*(2), 117–123. https://doi.org/10.1037/0003-066X.39.2.117