This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TAFFC.2018.2848656, IEEE Transactions on Affective Computing

An Affective Adaptation Model Explaining the Intensity-Duration Relationship of Emotion

John E. Steephen, Siva C. Obbineni, Sneha Kummetha and Raju S. Bapi, *Senior Member, IEEE*

Abstract—Intensity and duration are both pertinent aspects of an emotional experience, yet how they are related is unclear. Though stronger emotions usually last longer, sometimes they abate faster than the weaker ones. We present a quantitative model of affective adaptation, the process by which emotional responses to unchanging affective stimuli weaken with time, that addresses this intensity-duration problem. The model, described by three simple linear algebraic equations, assumes that the relationship between an affective stimulus and its experiencer can be broken down into three parameters. Self-relevance and explanation level combine multiplicatively to determine emotion intensity whereas the interaction of these with explanatory ease determines its duration. The model makes predictions, consistent with available empirical data, about emotion intensity, its duration, and adaptation speed for different scenarios. It predicts when the intensity-duration correlation is positive, negative or even absent, thus offering a solution to the intensity-duration problem. The model also addresses the shortcomings of past models of affective adaptation with its enhanced predictive power and by offering a more complete explanation to empirical observations that earlier models explain inadequately or fail to explain altogether. The model has potential applications in areas such as virtual reality training, games, human-computer interactions, and robotics.

Index Terms—affective adaptation, intensity and duration, modeling human emotion, region-β paradox

1 Introduction

The overall emotional experience of an event results from both the intensity and duration of the elicited emotion. Yet, while making daily choices, we may be tempted to focus a lot more on emotion intensity than how long it is likely to last, often leading to poorer decisions in life. The duration aspect has received only a limited attention even from cognitive scientists in spite of it being a pertinent feature of an emotional response. For instance, research that evaluates the efficacy of different stimuli in mood induction appears to focus only on the initial emotion strength and not how long it would last (e.g. [1]). While several studies and appraisal theories of emotion (e.g. [2], [3], [4]) have focused on its causes and elicitation, its duration has often been ignored [5], [6]. Studies that tried to elucidate the *relation* between the intensity and duration of an emotional response have been even fewer.

Different emotions have been reported to have different durations. For instance, it has been reported that joy and sadness generally last longer than fear and anger [7], [8]. It has also been reported that negative emotions have a longer duration than positive emotions [8], [9], [10], [11], [12]. It has been suggested (e.g. [5], [13]) and one would intuitively expect, that within an emotion type, the duration of an emotional experience would increase with its intensity. For instance, in a study by Gilbert, Lieberman, Morewedge, & Wilson [14], subjects expected the duration of their emotional response to correlate positively with its initial intensity.

There is some empirical evidence for this expected relation between intensity and duration. For instance, Sonnemans & Frijda [5] reported that the duration of an emotional response correlates with its intensity for positive emotions, disappointment, and sadness as long as there is no change in the situation that induced the emotion. The lengths of episodes of depressed mood also have been positively correlated with the severity of depression. The duration of the startle response has been reported to parallel stimulus strength, suggesting the same for the experienced fear underlying it [8]. A study by Rasinski, Berktold, Smith, & Albertson [15] showed that Americans who were more affected psychologically and emotionally by the 9/11 attacks retained the corresponding symptoms for a longer period on average. The amount of mental rumination and its duration have been observed to increase with initial emotion intensity [8], [16]. A prolonged duration of emotional consequences have also been found for extremely strong emotional events [8]. All these suggest a positive correlation between emotion intensity and its duration.

On the other hand, it has been reported that humans sometimes have a tendency to recover more quickly from deeply distressing incidents than the less painful ones [17], [18], [19]. Sonnemans & Frijda [5] observed a negative correlation between intensity and duration for fear. Rasinski et al. [15] found that Americans who knew someone killed or injured in the 9/11 attacks showed faster recovery from high stress levels as compared to those who did not. The experiments of Gilbert et al. [13], [14] provided additional evidence for this counter-intuitive phenomenon, who called this negative correlation the region-β paradox. Brandon & Silke [20] suggested that this paradox is perhaps present for anxiety problems as well, on the basis of the findings of

- *J.E. Steephen is with the Cognitive Science Lab, Kohli Centre on Intelligent Systems (KCIS), International Institute of Information Technology (IIIT), Hyderabad 500032, India. E-mail: johnsteephen@ gmail.com*
- *S.C. Obbineni is with the Centre for Neural and Cognitive Sciences, University of Hyderabad, Hyderabad 500046, India. E-mail: shivacogsci11@gmail.com*
- *S. Kummetha is with the Cognitive Science Lab, KCIS, IIIT, Hyderabad 500032, India. E-mail: kummetha.sneha@research.iiit.ac.in*
- *R.S. Bapi is with the School of Computer and Information Sciences, University of Hyderabad, Hyderabad 500046, India and with the Cognitive Science Lab, KCIS, IIIT, Hyderabad, 500032, India. E-mail: bapics@uohyd.ernet.in; raju.bapi@iiit.ac.in*

Smith, Perrin, Yule, & Rabe-Hesketh [21] and Thabet & Vostanis [22]. Thus it is clear that a simple relation does not exist between emotion intensity and its duration and the conflicting reports and arguments in the literature are hard to merge into a unifying hypothesis. We refer to this as the intensity-duration problem of emotion. The duration of an emotional response is primarily controlled by affective adaptation, the process by which active and passive psychological processes weaken emotional responses with time [6], [10], [11]. In addition to anecdotal accounts on affective adaptation, studies have demonstrated that we adapt to both pleasant (such as receiving an award or getting a promotion) and unpleasant (such as incarceration or losing money in a casino) incidents of life with time [10], [11]. The rate at which adaptation happens is known to differ across persons [12] and substantially between affective events [9], [12]. Three principles have been suggested to underlie affective adaptation on the basis of decades of research [6]. According to the *antagonism* principle, affective responses trigger conscious and subconscious processes that antagonize and weaken them. As per the *attention* principle, as subsequent events draw the attention from an emotional event, its emotional impact progressively decreases [6]. According to the *adaptation-level* principle or theory, emotional response decays as a result of the difference between stimulus strength and a reference point called adaptation-level, which is a function of past stimulus levels, reducing with time [6], [11], [23], [24].

In 2008, a simple affective adaptation model called AREA was proposed by Wilson & Gilbert [6]. According to AREA, if an affective event is self-relevant and poorly understood, it grabs our attention. Attending to it results in both an emotional reaction and an attempt at explaining the event. The process of explanation involves learning the nature of the event, determining its causes and understanding its consequences for our goals and self-concept. This results in the transformation of the event from being perceived as an extraordinary event to an ordinary one. As the event gets better and better understood by our attempts at explaining it, our attention to it also progressively reduces, resulting in our emotional reaction getting weaker with time. Once the event is completely explained successfully, it will no more hold our attention and we will not have an emotional reaction anymore. The process of adaptation is now complete. In this model, emotion intensity and its duration are determined by two factors, the self-relevance of the affective stimulus and how well or poorly understood the stimulus is to the experiencer. The greater the self-relevance and poorer the understanding, the stronger the emotion and longer its duration.

In spite of its elegance, AREA does not provide a satisfactory solution to the intensity-duration problem. To begin with, the model does not clearly specify the individual roles of self-relevance and understanding in determining intensity and duration. For instance, it does not predict the differences in emotional responses between two stimuli if one of them is more self-relevant but better understood than the other. The model also proposes that an event that is difficult to explain would enhance the emotional reaction. This is not fully consistent with empirical observations (see [25], [26]). For instance, in study 2 of Wilson et al. [26], there was no significant difference in the initial happiness intensity between the certain and uncertain conditions, though it was more difficult to explain the stimulus in the uncertain condition. This model also implicitly endorses a positive correlation between intensity and duration and does not accommodate the possibility of the region-β paradox. Further, it is based on the antagonism and attention principles and is reportedly inconsistent with the adaptation-level theory. Finally, it has not been formalized and cannot be used in artificial systems in the current form.

With respect to the intensity-duration problem, one pertinent question is whether duration can be predicted fairly reliably from intensity alone, i.e., is it a sole function of intensity or whether other factors also need to be considered. The former implies that the stimulus features that determine emotion intensity would also determine its duration. The latter would instead mean that the stimulus features determining intensity and duration are not identical. It needs to be stressed that an affective stimulus is a subjective quantity, and hence by stimulus feature we mean a feature of the stimulus-experiencer relationship and not any objective aspect of the stimulus. Though the examples given earlier suggest that intensity and duration are unlikely to share a monotonic relationship due to the absence of a positive correlation in some cases, the possibility of duration being a sole function of intensity through a non-monotonic relationship cannot be ruled out.

In 2013, a computational model of affective adaptation called HED was proposed by Steephen [27]. HED not only captured the primary features of affective adaptation but also made predictions consistent with several different characteristics of affective phenomena reported across literature. HED was based on the adaptation-level theory and it also proposed that adaptation rate is intensity dependent and that adaptation processes activate quickly but deactivate sluggishly. This resulted in an inverted-U shaped relation between intensity and duration, causing the duration to increase with intensity for lower intensities and decrease at higher intensities. This led to a positive correlation between them in some situations and the region-β paradox in others. HED argued, along the lines of the proposal of Gilbert et al. [14], that region-β paradox would occur only if an intense emotion is present. However, this is inconsistent with empirical data. For instance, the paradox was observed in the experiments of Gilbert et al. [14], where the stimulus, being just a mild attack on the self-esteem, would not have caused an intense emotion.

Since HED also aligned with the notion of duration being the sole function of intensity, it predicted emotional responses with the same initial intensity to have the same duration even if elicited by dissimilar affective stimuli. This, however, is not observed to be true, at least for positive emotions (see for e.g. [25], [26]). For instance, study 2 of Wilson et al. [26] demonstrated that in spite of equally happy emotions initially, the emotion lasted longer in the uncertain condition. This suggests that duration is unlikely to be an exclusive function of intensity and that the stimulus features that decide intensity and duration are not the same. Thus, while we feel that HED was an important step towards a comprehensive model of affective adaptation and emotion dynamics, and remains the only computational model to simulate and provide an explanation for the region-β paradox, it has also not addressed the intensity-duration problem sufficiently.

Yet the concepts underlying HED and AREA are based on years of affective adaptation research. In this paper, by adapting, modifying and extending these concepts, we propose a new computational model of affective adaptation that (1) addresses the limitations of past models, (2) is consistent with the antagonism, attention and adaptation-level principles and most importantly, (3) provides, perhaps for the first time, a plausible solution to the

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TAFFC.2018.2848656, IEEE Transactions on Affective Computing

intensity-duration problem consistent with empirical data.

2 Empirical Study

As discussed in the previous section, earlier research (e.g. [25], [26]) suggests that duration is not determined by intensity alone. Before developing the model, to strengthen this proposition, and to address the lacuna that this has not been shown for negative emotions, we conducted an empirical study. We used two video clips as stimuli that self-reports during pre-testing suggested to elicit similar sadness intensities. Subjects in the *fiction* condition watched the final scene of the movie *The Champ* [28], and those in the *non-fiction* condition watched an abridged version of the documentary *The Suicide Tourist* [29]. We expected subjects to adapt faster to the movie clip, it being merely a made-up story unlike the documentary clip which is an actual recorded event. Consequently, we hypothesized that subjects in both conditions would be comparably sad immediately after watching the clip, but that the subjects in the non-fiction condition would remain sad longer.

2.1 Method

Subjects. Ninety-one student volunteers participated in this study. The sample size was similar to that of an earlier study of the same type (study 2 of [26]).Written informed consent was taken before the experiment. During debriefing after the experiment, all of them reported that they had not previously watched their respective video clip. The data of subjects who reported neutral or happy emotions after watching the clip were excluded being irrelevant to the study. The resulting sample consisted of 42 subjects (22 m, 20 f) in the fiction condition and 37 subjects (17 m, 20 f) in the non-fiction condition.

Procedure. The experiment was conducted individually for each subject. The presentation of stimuli and collection of data were controlled using the PsychoPy software program [30]. Upon entering the laboratory, we explained the purpose of the experiment, the experiment procedure and its expected duration to the subjects. They read a short introduction to the clip they were assigned to and watched it on a 22” LCD monitor. The fiction and the non-fiction video clips were both around 8.5 min long. After watching their respective video clip, they performed the filler task of identifying emotions on 40 virtual faces that gave time for affective adaptation to happen. This was followed by a funny video clip to elevate their mood before debriefing and dismissing them. The subjective ratings of their emotional states made on a continuous rating scale whose end points were marked *very sad* and *very happy*, corresponding to the values -1 and +1 respectively, immediately after the clip and after the filler task were analyzed.

2.2 Results

A 2 x 2 mixed design ANOVA with one between-subjects factor (condition: fiction vs. non-fiction) and one within-subjects factor (time: right after clip vs. after filler task) revealed that the condition x time interaction was highly significant ($F(1, 77) = 12.17$, $p < 0.001$) consistent with our hypothesis (Fig. 1). The simple effect of condition was not significant right after the clip ($t(77) = -0.81$, $p = 0.42$, 2-tailed test; Cohen's d [95% CI] = -0.18 [-0.63, 0.27]), but the subjects in the non-fiction condition were significantly sadder than those in the fiction condition after the filler task ($t(77) = 2.68$, $p = 0.005$, 1-tailed test; $d = 0.61$[0.15, 1.06]), demonstrating that those in the non-fiction condition remained sad longer than those in the fiction condition in spite of comparable initial intensities. The experiment strengthened the conjecture that duration is not dependent on intensity alone and consequently that stimulus features that determine intensity and duration are not exactly the same. Secondly, the experiment showed that this conjecture is valid for negative emotions as well. Having gained sufficient confidence that duration is dependent more on specific stimulus features than on intensity, we moved on to model development.

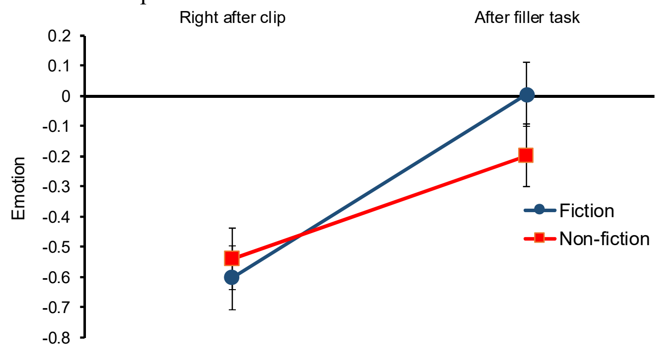


Fig. 1. Reported change in emotion intensity over time as a function of stimulus type. Means (with 95% CI) are ratings of how sad people felt. Higher negative numbers reflect greater sadness.

2 HED-ID (Human Emotion Dynamics – Intensity and Duration)

We put forward an affective adaptation model called HED-ID that proposes that emotion intensity and duration are determined by three parameters associated with the affective stimulus, namely *self-relevance*, *explanation level* and *explanatory ease*. Like AREA, we assume affective adaptation to happen through a process of explanation of the affective stimulus. We adapt the meaning of *explanation* from AREA and use it in an extended sense as follows. When an affective event occurs, our cognitive mechanisms start analyzing it so as to determine its nature, its meaning, how it fits into our self-concepts, its consequences for our goals, its causes and its coping potential. These include appraisal processes that estimate the self-relevance of the event by determining its implications for our goals and well-being [2], [3] as well as processes that help us to come to terms with it by reaching an understanding of it by determining its causes and addressing other knowledge gaps. Reaching an understanding refers to the assimilation of the event to one’s existing schemas or the alteration of these knowledge structures to account for the event [6]. We refer to the conscious and nonconscious processes involved in reaching an understanding of the event as explanation.

2.1 Self-relevance

Perceived self-relevance or desirability is one factor that determines emotion intensity as has been discussed in detail with ample evidence by Wilson et al. [6]. HED-ID also assumes initial emotion intensity to increase with the self-relevance of the event. According to AREA, elicited emotion ameliorates as people succeed in explaining the event or it stops being self-relevant. It may be emphasized that for emotion to dissipate, it is sufficient for either one of the above to happen. Thus, it is possible for the self-relevance of an event to remain unchanged, yet achieve emotion

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TAFFC.2018.2848656, IEEE Transactions on Affective Computing

abatement through successful explanation of the event. Alternatively, an event can remain poorly understood, yet the emotion can abate if it eventually turns out to be no more self-relevant.

Self-relevance can change and even disappear with time if the situation itself changes owing to the active steps taken by the person or due to other external factors. For instance, when the risk of death due to a certain disease gets greatly reduced by medical intervention, the disease becomes less self-relevant than before, bringing down the associated stress. Perceived self-relevance may also change on acquiring additional knowledge about or reassessing the event which is again nothing but a subjective change in the situation. After being pleased with topping a quiz, its self-relevance and the joy will reduce on learning that it will have no impact on the course grade. Similarly, when careful consideration reveals that the monetary loss from a business deal is not as big as initially thought, the event's self-relevance as well as the associated sadness decreases. A change in self-relevance is essentially a change in the affective stimulus itself and we take the position that emotion dissipation resulting from a change in the stimulus is not true affective adaptation. The abatement of sorrow over time even when a lost pet is not found is affective adaptation, whereas its abatement on finding the pet is not affective adaptation.

AREA model does not distinguish between cognitive processes that improve understanding and those that estimate or change the perceived self-relevance and refers to both as explanation. Consequently, emotion reduction resulting from a change in self-relevance also becomes affective adaptation in AREA. In contrast, along the lines of our argument that a change of self-relevance cannot be termed affective adaptation, we consider only those cognitive processes that improve the understanding of the stimulus as explanation.

It also needs to be noted that even after complete adaptation through a process of explanation, an event's self-relevance can remain unchanged. Consider someone getting a pay hike. They would be happy, but over time it gets explained and the happiness disappears. But, the raised pay is still self-relevant. Now if the hike is reversed, they would undoubtedly feel sad. This would not have happened had the hiked pay lost its self-relevance.

2.2 Explanation level

According to AREA, unexpected events attract attention and strengthen emotional responses since they are more difficult to explain than expected events, because when an event is expected, some explanatory work has already been done in advance. We do agree that for an expected event, the process of explanation would have started before the event occurrence, with the event being partially explained by the time it actually happens. Affective adaptation would also have thus started in anticipation of the event, a process that has been termed anticipatory adaptation [11]. Compared to an unexpected event, the amount of explanation to be done after event occurrence would be less in this case, allowing adaptation to complete faster. The *earlier* a future event is known about, the greater the amount of explanation completed by the time of event occurrence, and the shorter the time required to complete adaptation post the beginning of the event.

However, we do not agree with AREA that expectedness decreases the inherent *difficulty* in explaining an affective event. Instead, we believe that prior explanation only makes the event *better understood* by the time it happens. Some events are easy to explain, some are not, whether they are unexpected or not. An event can be completely unexpected and hence very poorly understood, but it may be very easy to explain. On the contrary, the expectation for an event may have begun a long time before its happening, causing the explanation process to start early, making it fairly understood by the time it happens. But it may be one that is difficult to explain that it may still take a long time for the explanation to complete after it has happened. For example, it is easier to explain the death of a pet cat if the cause is known as opposed to if it is unknown, irrespective of whether the death was expected or unexpected. Thus, an affective stimulus being poorly understood does not automatically imply that it is also hard to explain. Therefore, it is important to make a distinction between the extent of understanding of an event and the difficulty in explaining it and treat them as different factors.

This is in contrast to the approach of AREA where these two have not been unambiguously distinguished. In that model, the extent of understanding one has of an event is treated as the main determinant of the difficulty in explaining it and both these parameters have mostly been grouped together as a single entity that intensifies emotion and impedes adaptation. In HED-ID, we postulate that these two factors influence the emotional experience in different ways. While we agree with AREA that poor understanding intensifies the initial emotional reaction, we suggest that the difficulty in explaining the stimulus does not influence the initial emotion directly. Thus, the more unexpected an event, the more poorly understood it is, and the stronger the resulting initial emotion. In HED-ID, we refer to the degree of understanding of an affective stimulus at any given time as its *explanation level* and its value at the time of event occurrence as the *initial explanation level*. Two factors that determine the initial explanation level are the duration of expectation of a future event and the certainty with which the event is expected. The longer one expects the occurrence of a future event, or the greater the certainty with which one expects it, the lower the initial explanation level.

As per AREA, the stimulus attributes that aid or hinder explanation are novelty, surprise, variability, certainty, and explanatory coherence. We make a distinction between the first three variables and the remaining two. According to Wilson & Gilbert [6], a novel event is one which has not happened before, a surprising event is an unexpected experience, and a variable event is one which does not repeat in exactly the same way as before. Thus all three are essentially events which present themselves in an unexpected manner, that is, novelty, surprise, and variability are just different flavors of unexpectedness. When an unexpected event happens, it is novel, when an event happens at an unexpected time, it is surprising and when an event does not happen as expected, it is variable. Thus, novelty, surprise, and variability are the variables that decide how well or poorly understood an affective event is by controlling the extent of expectation.

2.3 Explanatory ease

Explanatory ease is the ease with which an event can be explained or adapted to. It is independent of whether the event is expected or unexpected. When the details of the event are *uncertain* or if it lacks *explanatory coherence*, the event becomes more difficult to explain. Thus certainty and explanatory coherence, the last two of the five variables suggested by Wilson & Gilbert [6] to aid or delay adaptation, are subservient to explanatory ease. Note that by explanatory coherence, we mean the factors other

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TAFFC.2018.2848656, IEEE Transactions on Affective Computing

than certainty that make a stimulus difficult or easy to explain and adapt to. As was noted by Wilson & Gilbert [6], an event may affect many people in the same way, but the difficulty in explaining it, i.e. explanatory ease, can be different for each of them depending on the methods of explanation chosen.

Like AREA, HED-ID proposes that explanatory ease influences the rate of explanation but unlike AREA, does not propose that it affects initial intensity directly. However, it can affect initial intensity indirectly for anticipated events. As stated earlier, anticipating an affective event can make it better understood by the time it happens. How much better understood it becomes depends on the amount of explanation completed during the anticipation phase, which in turn is determined by the ease with which it can be explained. Thus the explanation level at the time of the event depends on, in addition to how much it is expected and how early it is anticipated, how easy it is to explain.

2.4 Explanatory effort

There is one more factor other than explanatory ease that influences the rate of explanation and that is the amount of cognitive resources allocated to explain the event. For an event with a given explanatory ease, the more the cognitive resources allotted, the faster one will be able to explain it.

To explain the region-β paradox, Gilbert et al. [14] had proposed that a strong emotion leads to the rapid mobilization of several adaptation processes which speeds up adaptation, causing the emotion to die down faster than other less intense emotional responses. They suggested that these psychological processes get initiated *only* when the intensity of emotion goes above a certain critical threshold. Since adaptation always happens irrespective of emotion intensity, this conjecture raises the question as to how one would adapt to weaker emotions whose intensity does not cross that threshold. Hence, this proposal was modified by Steephen [27] who suggested that it is more plausible for adaptation processes to increase monotonically in strength and number with emotion magnitude than the existence of a single threshold when all the processes get switched on suddenly.

Adapting this idea to HED-ID, which assumes explanation as the process underlying adaptation, we propose that an event eliciting a strong emotion would result in the exertion of a large cognitive effort, in terms of the number and strength of the cognitive mechanisms assigned, to explain it. On the contrary, a weak emotion would cause only a minimal effort being put towards explaining the event. Thus, through a greater cognitive effort, a larger emotion would help increase the speed of explanation. Care must be taken to note that in this paper, the term explanatory

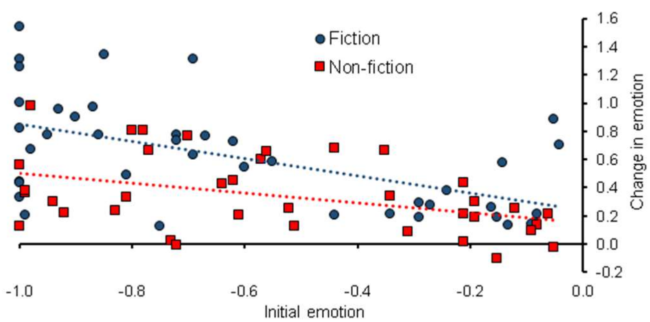


Fig. 2. Reported change in emotion intensity as a function of initial emotion intensity across subjects. Change in emotion has a significant correlation with initial emotion intensity under both fiction and non-fiction conditions. The dotted lines show the results of linear regression.

effort (or adaptation effort) refers to the amount of *effort put in* and not the amount of *effort required* to explain (or adapt).

This proposition is also supported by Rimé et al. [16] who showed that adaptation processes such as mental rumination and social sharing increases with emotion intensity. Further, this proposal is consistent with AREA which associates a greater emotion with greater attention. One would expect that paying more attention to an event would lead to an increase in the effort exerted to explain it. Additional evidence may also be present in the data from the empirical study in this paper. Fig. 2 shows the relationship between the initial emotion intensity and the change in emotion over time for each subject in the fiction and non-fiction conditions. Linear regression analysis suggests that across subjects, the reduction in emotion correlates positively with initial intensity in both the fiction ($r = 0.56$, $p < 0.001$) and non-fiction ($r = 0.42$, $p < 0.01$) conditions. That is, within each condition, the subjects who reported a higher initial sadness were also the ones who recovered faster during the filler task. This is indicative of and consistent with our proposal since this is what would happen if the explanatory effort increased with intensity.

2.5 Quantitative specification

According to AREA, an event triggers an affective reaction only when it is both self-relevant and unexplained. Adaptation is complete when either the explanation is complete or when it is no longer self-relevant. Further, the elicited affective reaction has been shown to increase with self-relevance and decrease with the understanding of the event. On the basis of this, one could assume that self-relevance (SR) and explanation level (EL_t) of an event at a given time (t) combine multiplicatively as follows to determine the magnitude of the emotion (u_t) at that time.

$$u_t = SR \times (1 - EL_t) \quad (1)$$

A value of 0 for explanation level indicates that the event is completely unexplained, whereas a value of 1 means that it has been fully explained. The highest possible emotion intensity for a given self-relevance will happen when explanation level is 0. When an event occurs totally unexpectedly (surprise), or when a completely unexpected event (novel) occurs, explanation level will be 0 initially. An affective reaction proportional to the self-relevance of the stimulus will result. As one goes through the psychological process of explaining the event and it becomes more understood, its explanation level slowly rises and approaches 1. Consequently, the emotion also abates with time and disappears completely when explanation level reaches 1. If an affective stimulus has been expected since a long time, and it happens in exactly the same way and at the same time as anticipated, it would have been wholly explained by the time the event actually happens. There would be nothing more to explain and the initial explanation level will be 1. So even if it is high in self-relevance, there will be no emotional response when it happens.

However, in real life, an initial explanation level of 0 or 1 is unlikely since hardly anything ever happens in an entirely expected or unexpected way. The degree of expectedness just differs between events. Further, this expectedness can begin at any time before the event, with the time elapsed until the event influencing the initial explanation level. If an event is only partially surprising or novel, or different from previous occurrences, i.e., if the event is partially unexpected (the event itself or its occurrence), it would also be a partially explained event. When it occurs, its explanation level would be between 0 and 1 and depend on how much more explanation remains to be done. The initial emotion would be proportional to its self-relevance, but discounted by the explanation level.

As mentioned earlier, the explanation rate at a given time is dependent on the explanatory ease of the event and the explanatory effort exerted. We assume a linear relationship between emotion intensity and explanatory effort, and consequently a similar relationship between emotion intensity and the explanation rate (r_t). The slope of this relationship would depend on how simple or hard it is to explain the stimulus and hence it is determined by the explanatory ease. Explanatory ease (*EE*), being a characteristic of the event and its experiencer is considered a constant like self-relevance and assumed to not change with time.

$$r_t = EE \times |u_t| \quad (2)$$

Explanation rate is a time varying quantity and it refers to the amount of explanation done per unit time. At any given time instant, to find the explanation level for the next instant, the current explanation level has to be added with the amount of explanation that will be done between the current and next instants. The amount of explanation that will be done during this time is the explanation rate at the current instant. Thus, once the rate at which explanation is proceeding has been determined for a given time instant (t) using (2), the explanation level at the next instant ($t+1$) will be given by

$$EL_{t+1} = EL_t + r_t \quad (3)$$

The new explanation level along with self-relevance will determine the new emotion intensity. This will continue until the emotion decays completely. The process of affective adaptation according to HED-ID is shown in Fig. 3. The Python programming language was used to implement HED-ID computationally and run simulations. The executable model code is provided as supplemental material. Affective stimulus parameters can be varied and simulations run using its graphical user interface.

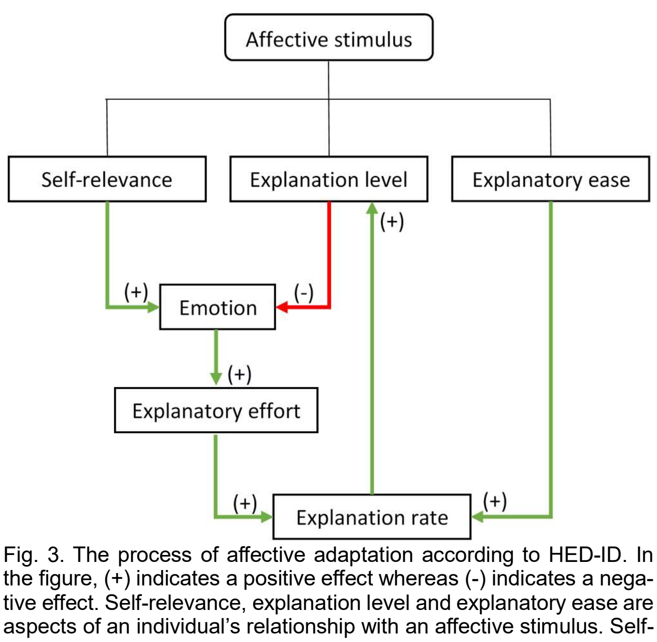


Fig. 3. The process of affective adaptation according to HED-ID. In the figure, (+) indicates a positive effect whereas (-) indicates a negative effect. Self-relevance, explanation level and explanatory ease are aspects of an individual's relationship with an affective stimulus. Self-relevance and explanation level together determine emotion intensity where self-relevance has a positive effect and explanation level a negative effect. The rate at which explanation happens increases with both emotion intensity, through increased explanatory effort, and explanatory ease of the stimulus. The rate at which explanation level increases is decided by this explanation rate. The increasing explanation level decreases emotion intensity. The process continues until emotion decays completely.

3 Model Predictions

HED-ID is a quantitative model that assumes a certain magnitude for all the variables it deals with. Their counterparts in the real world also have magnitudes and are quantitative entities. For instance, there are magnitudes for the perceived self-relevance of an event, the extent of understanding one has of an event, the intensity of emotion elicited, the amount of attention paid to an event, the amount of active and passive cognitive efforts put in to explain an event and so on. However, these are primarily subjective quantities, and due to limitations in the accurate quantitative measurement of subjective phenomena, a quantitative comparison between the model predictions and real life observations is difficult. Nevertheless, to examine whether the model is consistent with empirical data, one can look at scenarios where an ordinal comparison between multiple affective stimuli as well as their corresponding emotional responses is possible.

Now we discuss the predictions made by HED-ID under nine different scenarios using computer simulations. For each scenario, we consider two affective stimuli designated X and Y that differ in one or two parameters. X and Y can be considered as two non-overlapping events happening in an individual's life or as events happening to different individuals or groups of individuals. They are never two events happening simultaneously in a single person's life. We make ordinal predictions of the emotional responses using HED-ID. We also compare them with the findings from every instance of the scenario we could find in empirical literature.

The emotion response characteristics we are interested in are emotion intensity, its duration and overall adaptation rate. Emotion duration is the time taken for an elicited emotion to completely vanish or reach a negligible level. For the purpose of our simulations, we take this level as 1% of the maximum possible intensity. The overall adaptation rate is the average rate at which emotion decays from the beginning to the end of adaptation. It is the ratio of initial emotion intensity to duration. Table 1 summarizes the results from our simulations and Fig. 4 shows representative simulation outputs for each scenario.

Some predictions such as those for Scenarios 1, 2 and 4 (see Table 1) may not be surprising, given the assumptions of the model, whereas the others are not so obvious. The predictions for Scenario 1 are interesting in that, unlike the predictions of prior models such as AREA and HED, they underscore the principle that *intensity is not the sole predictor of duration.* Scenario 1 happens when two events differ in certainty or explanatory coherence. In Kurtz et al. [25] and studies 1A, 1B, 2 and 3 of Wilson et al. [26], when certainty alone was manipulated, the initial intensity was unaffected, but a slower adaptation and a longer duration for the uncertain condition resulted. As according to HED-ID, certainty manipulation would affect explanatory ease but not self-relevance or initial explanation level, these experiments are likely instances of Scenario 1. As is evident from Table 1, HED-ID's predictions are in accordance with the experimental observations detailed below.

In Kurtz et al. [25], subjects were told that they would receive gifts. Those in the certain condition were told immediately which gifts they would receive whereas those in the uncertain condition were not given this information until the end of the study. The initial happiness was similar for both the groups but the subjects in the uncertain condition remained happy longer.

In study 1A of Wilson et al. [26], students in a library were

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TAFFC.2018.2848656, IEEE Transactions on Affective Computing

TABLE 1
SUMMARY OF RESULTS FOR DIFFERENT SCENARIOS

Scenario	*SR*	*EL*	*EE*	Emotion intensity	Emotion duration	Overall adaptation rate	Intensity-duration relation	Consistent empirical observations
1	X=Y	X=Y	X>Y	X=Y	X<Y	X>Y	EE dependent	Studies 1A, 1B, 2, 3 in [26]; Study 1 in [25]; Study 4 in [31]
2	X=Y	X<Y	X=Y	X>Y	X>Y	X>Y	Positive correlation	[9]
3	X>Y	X=Y	X=Y	X>Y	X<Y	X>Y	Region-β	[5]; [15]; Study 2, 3 in [14]
4	X=Y	X<Y	X<Y	X>Y	X>Y	EE, EL dependent	Positive correlation	Study 6 in [31].
5	X=Y	X<Y	X>Y	X>Y	X<Y	X>Y	Region-β	NA
6	X>Y	X=Y	X>Y	X>Y	X<Y	X>Y	Region-β	NA
7	X>Y	X=Y	X<Y	X>Y	SR, EE dependent	SR, EE dependent	SR, EE dependent	NA
8	X>Y	X<Y	X=Y	X>Y	X<Y	X>Y	Region-β	NA
9	X>Y	X>Y	X=Y	X=Y	Y>X	X>Y	No correlation	NA
				X<Y	Y>X	SR, EL dependent	Positive correlation	NA
				X>Y	Y>X	X>Y	Region-β	NA

SR = Self-relevance, EL=Explanation level, EE = Explanatory ease

gifted a card with a United States golden dollar coin attached. In the uncertain condition, there was ambiguity about the source of the gift whereas there was no ambiguity in the certain condition. Similarly, in study 1B of Wilson et al. [26], students in a cafeteria were gifted a card with a golden dollar coin and a dime. In the uncertain condition, there was ambiguity about the odd amount whereas there was no such ambiguity in the certain condition. Around 5 minutes later in both these studies, the students in the uncertain condition were happier, implying that the emotion elicited by the gift lasted longer in the uncertain condition.

In study 2 of Wilson et al. [26], participants watched an upbeat film based on the true story of a man named Rudy. Participants in the certain condition were unambiguously informed what happened to Rudy after what was shown in the film. Participants in the uncertain condition were given two accounts of what happened to him and were told that only one of them was true. It was observed that although positive emotions of similar intensity were elicited in both the groups, the emotion lasted longer in the participants in the uncertain condition.

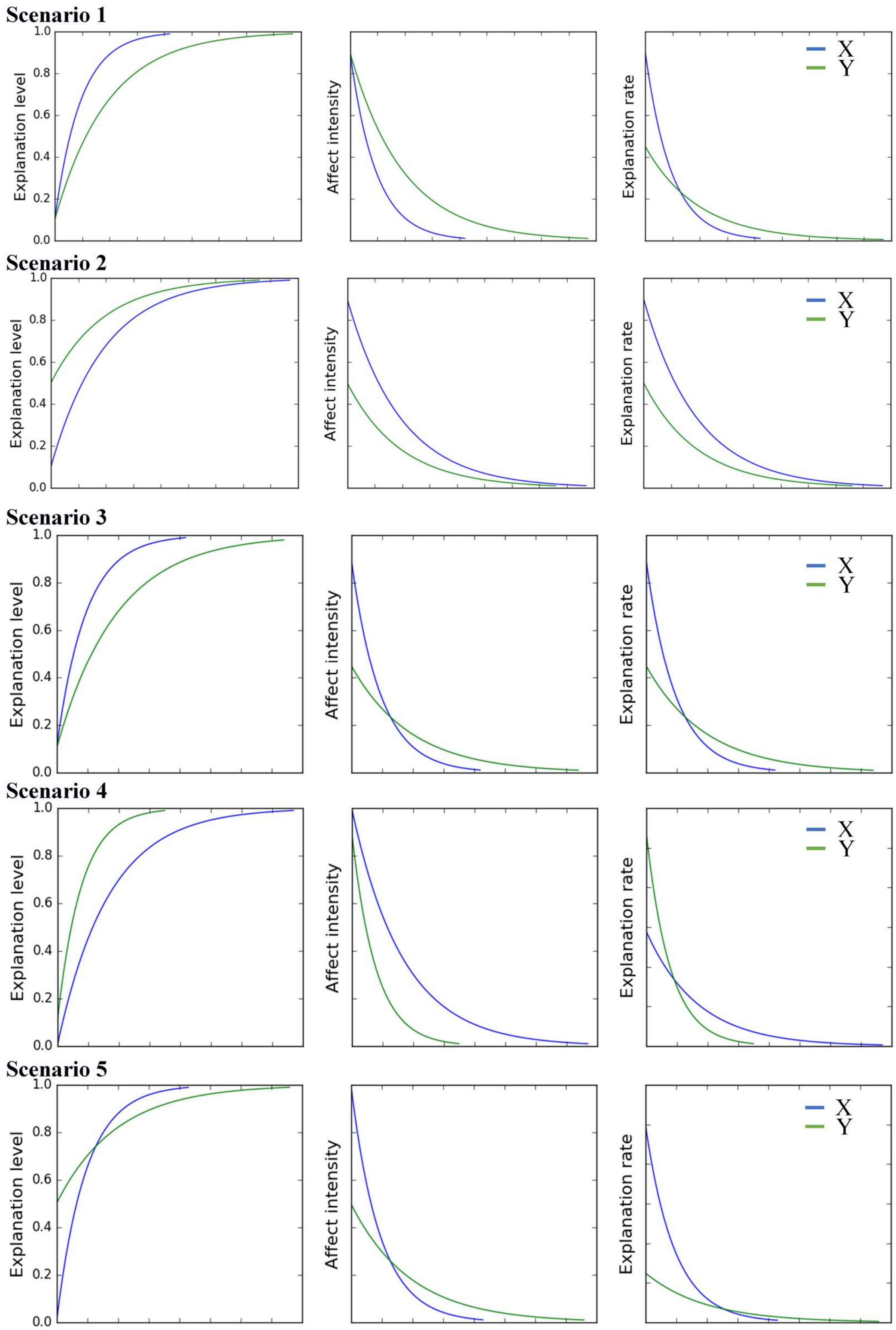
In study 3 of Wilson et al. [26], student participants were made to believe that they received positive feedbacks from three opposite-gender students. In the certain condition, the participants got to know which feedback was written by each student whereas in the uncertain condition, they only got to know who the three students were. Though the initial emotion intensities of the participants in both the conditions were not statistically different, those in the uncertain condition maintained their positive emotion longer.

Study 4 of Gilbert et al. [31] is another potential instance of Scenario 1, with explanatory coherence the likely manipulated variable. In this study, the personalities of subjects with positive and negative self-views were rated poorly. Those in the fallible and infallible conditions were then told that they were assessed by a computer and a team of experts respectively. The self-relevance of the undesirable classification and the initial explanation level would be the same for the subjects with a positive self-view in both the conditions. The explanatory ease would be greater for those in the fallible condition as it is easier to explain an undesirable classification from a potentially incompetent computer program as opposed to a team of experts. In this scenario, HED-ID would predict similar initial emotion intensities for both the groups and a longer emotion duration for those in the infallible condition. This is consistent with the experiment finding that those in the infallible condition felt worse after 5 min, implying a longer emotional response for them. No significant difference was observed in the emotion after 5 min between subjects with a negative self-view in the fallible and infallible conditions. Since their initial emotion was not assessed, it is not known whether they were even affected by the stimulus. It is possible that they expected the undesirable classification due to their negative self-view. This self-view may have caused anticipatory adaptation to any stimuli that could hurt one's self-esteem and consequently could have prevented any emotional response. It may be noted that unlike HED-ID, both AREA and HED are inconsistent with the results of any experiment matching Scenario 1.

On the basis of a meta-analysis of close to 200 studies, Luhmann et al. [9] suggested that marriage, childbirth, and divorce are more likely to be actively initiated than bereavement and unemployment. Hence, the former are events one starts expecting months or years early. Indeed, one anticipates the birth of a child several months in advance. Therefore, the initial explanation levels of the former would be higher. This can be considered an instance of Scenario 2, if we were to assume the differences in self-relevance and explanatory ease between the two event categories to be insignificant compared to that of explanation level. Consistent with HED-ID's prediction, they found stronger affective reactions for the latter. The logarithmic adaptation rate was reported to be similar for all the events. If two stimuli with different initial intensities have the same logarithmic adaptation rate, the one with the higher intensity will have the longer duration. This suggests that bereavement and unemployment have a longer emotion duration, again consistent with the model's predictions.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TAFFC.2018.2848656, IEEE Transactions on Affective Computing

In study 6 of Gilbert et al. [31], subjects were told they would be selected for a job on the basis of a screening procedure that involved answering a set of questions. Subjects in the unfair condition received questions that appeared irrelevant to the hiring decision and were told that the decision would be made by one expert. Those in the fair condition received questions that appeared a lot more relevant and were told that they would be rejected only if three experts unanimously decided so. After evaluation, each of them was told that they were not selected. The self-relevance of rejection would be the same for all subjects. It would be easier to explain rejection by one expert on the basis of answers to irrelevant questions than rejection by all three experts on the basis of answers to relevant questions. Since all the subjects would have considered the possibility of rejection from the beginning itself, explanation would have started in anticipation much before the rejection. The explanation level at the time of rejection would be higher for those in the unfair condition due to two reasons. First, they would expect to get rejected more than



This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TAFFC.2018.2848656, IEEE Transactions on Affective Computing

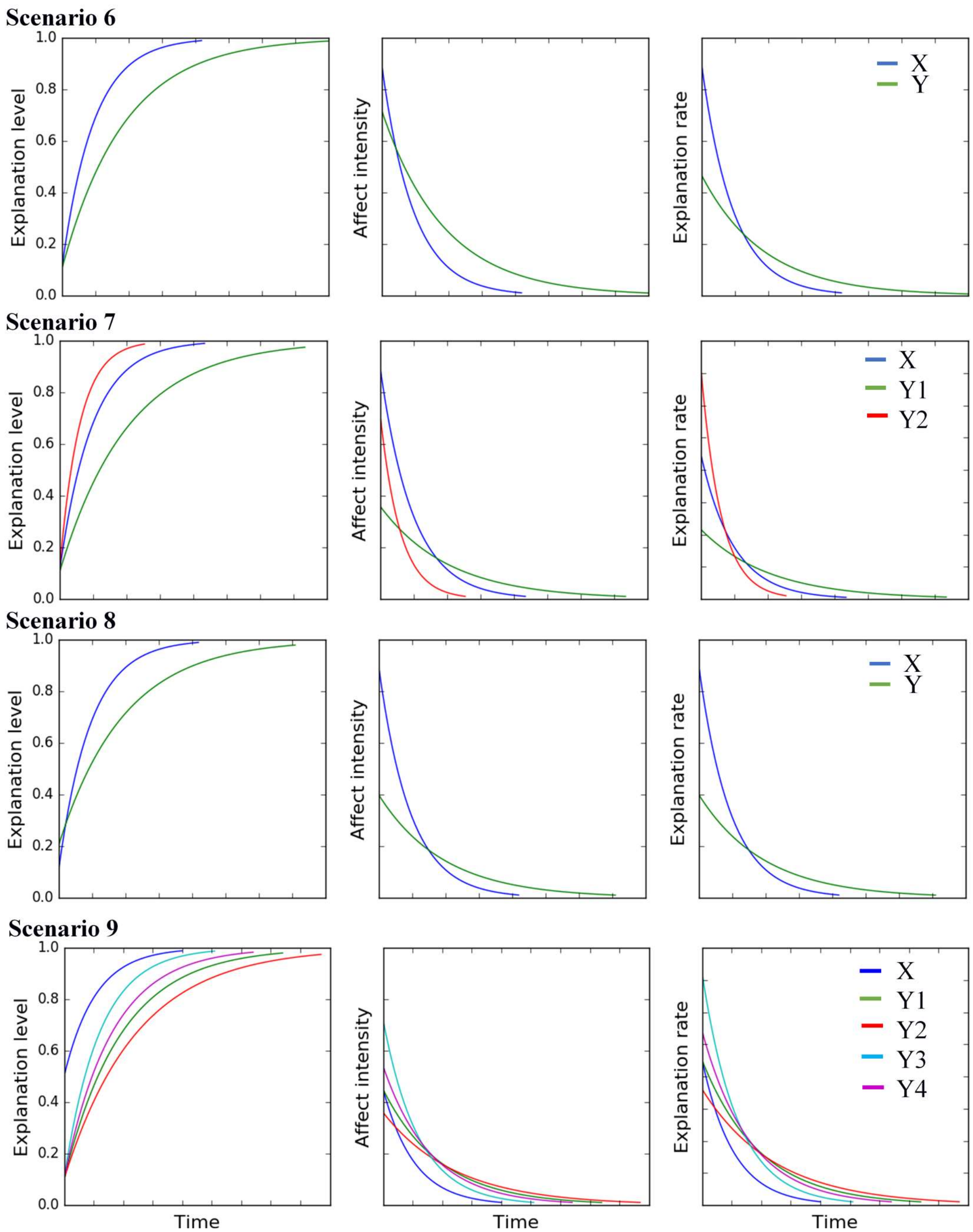


Fig. 4. HED-ID's predicted explanation and emotion dynamics for two affective stimuli, X and Y under different scenarios. In *Scenario1,* X has the same self-relevance and is as poorly understood initially as Y, but is easier to explain. X and Y will elicit the same initial intensity but X's adaptation completes faster due to its greater explanatory ease. In *Scenario 2*, X has the same self-relevance and is as easy to explain as Y, but is less understood. X elicits a higher initial intensity than Y and maintains the higher relative intensity throughout adaptation, resulting in its emotion to last longer. In *Scenario 3*, X is equally understood and as hard to explain as Y, but more self-relevant. X elicits a higher initial intensity than Y due to its higher self-relevance, but in due course, its emotion dies down faster thanks to quicker adaptation, demonstrating the region-β paradox. In *Scenario 4*, X is as self-relevant as Y but less understood and harder to explain. X elicits a stronger emotion than Y due to its poorer understanding and the intensity remains higher throughout, also causing the emotion to last longer. In *Scenario 5*, X is as self-relevant as Y but less understood and easier to explain. X elicits a stronger emotion than Y due to its poorer understanding but the emotion drops below that of Y in due course resulting in a shorter emotion duration and exhibiting the region-β paradox. In *Scenario 6*, X is as understood as Y but more self-relevant and easier to explain. X elicits a stronger emotion due to its higher self-relevance but the emotion drops below that of Y in due course, resulting in a shorter emotion duration and exhibiting the region-β paradox. In *Scenario 7*, X is as understood as Y but more self-relevant and harder to explain. Y elicits a weaker emotion due to its lower self-relevance, but the emotion may (Y1) or may not (Y2) go above that of X in due course depending on the relative values of self-relevance and explanatory ease. Consequently, Y's emotion duration may be longer (Y1) or shorter (Y2) than that of X and thus the region-β paradox may (X & Y1) or may not (X & Y2) occur. In *Scenario 8*, X is as hard to explain as Y, but is less understood and more self-relevant. X elicits a stronger emotion due to its higher self-relevance and lower understanding but the emotion will drop below that of Y in due course causing the region-β paradox. In *Scenario 9*, X is as hard to explain as Y, but is more self-relevant and understood. X can elicit an emotion as strong as (Y1), stronger than (Y2) or weaker than Y (Y3 & Y4), but it will always have the shorter duration. Hence, if the initial emotion intensity of Y is less than that of X, region-β paradox can be observed (compare X & Y2).

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TAFFC.2018.2848656, IEEE Transactions on Affective Computing

those in the fair condition since they were being evaluated by only a single individual. This is because, if we take the chance of getting rejected by a single judge as 50%, the chance of getting rejected by all the three judges would be only 12.5%. A higher expectation would increase the initial explanation level. Secondly, anticipatory explanation would also have been easier for these subjects because, as they were going through the experiment, they would feel the unfairness in the evaluation process and could easily explain that if they did not get selected, it would be because of this unfairness and not because of their incompetence. The easier anticipatory explanation would also raise the initial explanation level. This is an instance of Scenario 4 and consistent with HED-ID's predictions, subjects in the fair condition felt worse than those in the unfair condition immediately after rejection and also their emotion lasted longer. Further, the authors reported the difference in emotion intensity between the subjects of the two conditions to widen with time, which as shown in Fig. 4, is also consistent with HED-ID.

The prediction of the region-β paradox in Scenarios 3, 5, 6 and 8, which cannot be easily deduced from the underlying principles, nonetheless arises from them. The paradox's prediction in Scenario 3 is particularly intriguing. Consider the study by Rasinski et al. [15] on the psychological and emotional symptoms experienced by Americans after the 9/11 attacks. Naturally the event was more self-relevant to those who knew someone hurt or killed in the attacks than those who did not. The event was equally unexpected by both the groups and hence their initial explanation level was the same. Both the groups were trying to explain the same event, and hence the explanatory ease would be similar for both. These conditions represent Scenario 3 and HED-ID would predict the more affected group to have a higher initial emotional intensity, which is consistent with the authors' report that they had more psychological and emotional symptoms initially (5.5 vs. 4.3). HED-ID would also predict that this group would adapt faster in that they would have a higher overall adaptation rate as well as a shorter emotion duration. In other words, HED-ID would predict the region-β paradox. When the symptoms of both the groups were evaluated after 4 to 6 months, surprisingly this is exactly what was found. In the group that was affected more, the symptoms reduced by 3 from 5.5 to 2.5, whereas the symptoms of the less affected group reduced by only 1.7 from 4.3 to 2.6. A similar trend was seen even when the population of New York alone was considered. Also, the initial reaction for those from the city was stronger than those from all over America, since the event has greater self-relevance for them, New York being the site of attack.

In study 2 of Gilbert et al. [14], the personalities of the participants were rated poorly. They were either told that they would meet their rater (partner condition) or that they would not (non-partner condition). One would expect the negative rating to be more self-relevant to those who have been told that they would meet their rater. All the participants would have the same initial explanation level as the affective stimulus was equally unanticipated. The explanatory ease would be the same for participants in both the conditions since what they have to explain is the same. Thus this would be another instance of Scenario 3. Region-β paradox, as predicted by HED-ID, was observed when the participants' responses were assessed immediately after the rating and 5 min later.

In study 3 of Gilbert et al. [14], the participants were divided into the victim and bystander groups. The procedure for the victims was identical to those in the partner condition of study 2. Bystanders were provided with all the information the victim had including that the victim was rated poorly by the rater. They were also told that they would later interact with the victim but not the rater. After 5 min, the bystanders were found to dislike the rater more than the victims. Obviously the victims would have been emotionally more affected initially as the negative rating was more relevant to them. The initial explanation level was the same for both the bystanders and the victims because the affective stimulus was equally unexpected by both of them. There is no reason to assume a different explanatory ease for both the groups. Thus this would again be an instance of Scenario 3 and HED-ID would predict the victims to adapt faster causing them to dislike the rater less than the bystanders after some time, consistent with the empirical observation.

Sonnemans & Frijda [5] had reported a negative correlation between intensity and duration in the case of fear. As compared to emotions such as happiness and sadness, fear is not usually anticipated. Thus incidents that cause fear may be assumed to have comparable initial explanation levels. This suggests that the different initial fear intensities from different events are likely to be more due to differences in self-relevance than initial explanation levels. If the explanatory ease of these events are also not so different, it is a case of Scenario 3, and one can expect a negative correlation between intensity and duration.

Unlike other scenarios, Scenarios 7 and 9 have multiple possible outcomes. For Scenario 7, the emotional response to the more self-relevant stimulus is stronger initially. The stronger emotion would lead to more cognitive resources being allotted for explanation, but since the stimulus is harder to explain, the rate at which explanation happens can be greater or less than the less self-relevant stimulus. Consequently, either stimulus can get explained faster depending on their specific values of self-relevance and explanatory ease. In contrast, in Scenario 9, while the emotional response of the more self-relevant stimulus can be stronger than, weaker than or equal to that of the less self-relevant stimulus depending on the specific values of self-relevance and initial explanation level of the stimuli, the latter would always last longer. Empirical instances matching Scenarios 5 to 9 could not be located in literature.

3 HED-ID and Adaptation-Level Theory

Since some of the core ideas underlying HED-ID are adapted from AREA, it is consistent with the antagonism and attention principles. However, unlike AREA, we now show mathematically that HED-ID is within the framework of adaptation-level theory also. According to adaptation-level theory [11], [23], [24], the emotional response to an affective stimulus is determined by the difference between the stimulus strength and a reference point called the adaptation-level. Adaptation-level is a function of past stimulus levels. At the beginning of the stimulus when adaptation has not yet begun, adaptation-level is zero and the emotion elicited is a function of the stimulus intensity alone. With time, the adaptation-level increases since it is determined by the previous stimulus levels. This causes the elicited emotion to decrease with time as the gap between the stimulus strength and the adaptation-level reduces, leading to affective adaptation. When the adaptation-level reaches the stimulus level, that is,

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TAFFC.2018.2848656, IEEE Transactions on Affective Computing

when it becomes equal to the strength of the stimulus, adaptation is complete and no more emotion results. Thus, adaptation-level is nothing but a measure of how much one has adapted to an affective stimulus. We suggest that it is the process of explanation that causes adaptation-level to increase as a function of past stimulus levels, that stimulus intensity corresponds to the self-relevance of the stimulus and that the rate at which adaptation-level rises is determined by the self-relevance and explanatory ease of the stimulus.

Adaptation-level is often formulated as a weighted average of past stimulus levels with the weight decreasing towards the older stimulus levels [11], [27], [32] using an equation such as:

$$AL_t = \alpha X_{t-1} + (1 - \alpha)\, AL_{t-1} \tag{4}$$

where AL and X are the adaptation and stimulus levels respectively and α represents the speed of adaptation. If we consider the stimulus level to not change with time, like the affective stimuli we dealt with in this paper before, X would be a time invariant quantity and (4) can be written as:

$$AL_t = \alpha X + (1 - \alpha)\, AL_{t-1} \tag{5}$$

Using (1) – (3), the formula for explanation level can be rewritten as (the derivation is provided as appendix):

$$EL_t = SR \times EE + (1 - SR \times EE)\, EL_t \tag{6}$$

How the adaptation-level starts from zero and changes over time until it reaches the stimulus level is given by (5), whereas how the explanation level starts from zero and approaches 1 is given by (6). It may be noted that both the equations are identical in their form. Adaptation-level theory does not state why adaptation-level changes as a function of past stimulus levels. The theory also does not specify the factors that determine the speed of affective adaptation, α.

Comparing (5) and (6), we suggest that it is the process of explanation that causes the adaptation-level to rise over time and therefore it essentially mirrors the explanation level. We suggest that the adaptation-level is determined by the explanation and stimulus levels and that the adaptation speed in the adaptation-level model is determined by the self-relevance and explanatory ease of the stimulus as follows:

$$AL_t = EL_t \times X \tag{7}$$

$$\alpha = SR \times EE \tag{8}$$

It may be noted from (7) that when the explanation level is zero, the adaptation-level is also zero, and when the explanation level reaches 1, the adaptation-level reaches the stimulus level.

According to adaptation-level theory, emotion intensity is a function of the difference between the stimulus and adaptation-levels, i.e.

$$u_t = \mathrm{f}(X - AL_t) \tag{9}$$

Unlike a physical stimulus, an affective stimulus does not have an objective intensity. One way to deal with this problem is to specify it in terms of the intensity of the emotion it elicits before adaptation sets in. This approach can be taken since stimulus intensity is nothing but a measure of its desirability and the emotion before the onset of adaptation would correspond to this desirability (See [27]). Thus a stimulus that would elicit 10 units of emotion can be taken to have a strength of 10 units. Using this approach, (9) can be simplified as:

$$u_t = X - AL_t \tag{10}$$

According to HED-ID, emotion intensity is determined by a multiplicative combination of self-relevance and explanation level as given by (1). The equation can be rewritten as:

$$u_t = SR - SR \times EL_t \tag{11}$$

From this equation, it can be seen that the emotion intensity before the onset of adaptation is given by the self-relevance of the stimulus since the explanation level is zero at that point. In the adaptation-level model, this emotion intensity is given by X (see (10)). Replacing SR with X in (11),

$$u_t = X - X \times EL_t \tag{12}$$

Now since $AL_t = (X \times EL_t)$ as shown in (7), (12) becomes

$$u_t = X - AL_t \tag{13}$$

which is the equation for emotion intensity in the adaptation-level model. Thus, starting from the equation for emotion intensity in HED-ID, we arrive at the same in the adaptation-level model. Thus we show that HED-ID is consistent with the formulation for the adaptation-level theory and suggest that the stimulus intensity, the adaptation-level and the adaptation speed of the adaptation-level theory are given by the self-relevance, the product of self-relevance and explanation level, and the product of self-relevance and explanatory ease of HED-ID respectively.

4 Discussion

Using a new computational model of affective adaptation called HED-ID, we have offered a plausible solution to the intensity-duration problem of emotion. In this model, the relationship between an individual or a group and an affective event was decomposed into three parameters namely, self-relevance, explanation level, and explanatory ease. Persuaded by past and our own empirical observations, the model adheres to the principle that emotion intensity and its duration are determined by different stimulus feature sets. Thus, self-relevance and explanation level determine intensity, whereas the interaction of explanatory ease with them decides duration. Our model shows why intensity correlates with duration sometimes, why stronger emotions die faster on some occasions, and how similarly intense emotions can have a range of durations. For instance, if the difference in emotion intensity between two stimuli is due to one of them being anticipated, the stronger response would take longer than the weaker one to subside (Scenario 2), whereas if the difference is due to a difference in self-relevance, the stronger emotion would die down faster (Scenario 3). HED-ID is also consistent with the three principles that have been proposed to explain affective adaptation namely the attention principle, the antagonism principle, and the adaptation-level theory (see [6]).

4.1 Affective adaptation curve

One thing of interest is how the emotion magnitude changes over time during the process of affective adaptation. In HED-ID, emotional responses decay as a negatively accelerated curve, that is, the change in emotion intensity per unit time decreases with time (see Fig. 4). Though it has not been empirically tested if emotion dissipation is better approximated by a linear or negatively accelerated curve, there is more evidence suggesting the latter as has been argued by Steephen [27]. The observation of Luhmann et al. [9] that logarithmic change models have a better fit than linear change models for the decay of emotions from major life events and the apparent negatively accelerated dissipation of the distress induced by the fear of contracting an incurable disease [33] further support that argument. Even the data from our experiment shown in Fig. 2 is more consistent with a negatively accelerated course of emotion amelioration as it suggests emotion change to be proportional to its initial intensity. This pattern

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TAFFC.2018.2848656, IEEE Transactions on Affective Computing

of decay can be explained by the intensity dependence of explanatory effort that we have proposed.

4.2 AREA, HED and HED-ID

AREA and HED are both affective adaptation models with the former incorporating the antagonism and attention principles and the latter based on the adaptation-level theory. Though HED-ID shares some features with these models, there are aspects in which it differs from either one or both of them. For instance, whereas in HED, adaptation speed was treated as more of an individual's property and less of the nature of the event, HED-ID, along the lines of AREA, views it as a property of the relationship between the individual and the event in accordance with the suggestion that it varies across persons and situations [12]. On the other hand, similar to HED but unlike AREA, HED-ID proposes that the effort put into adaptation increases with emotion intensity. HED and HED-ID predict the time course of emotion during the process of adaptation and suggest it to follow a negatively exponential curve whereas the AREA model does not discuss the shape of the curve at all. HED and HED-ID explain the region-β paradox, whereas the AREA model does not. However, the mechanisms proposed by HED as the basis of the paradox are different from what HED-ID suggests. HED suggests that the paradox results from the combined effect of an emotion-dependent adaptation speed and sluggish deactivation of adaptation processes. Instead HED-ID shows how differences in explanatory ease, initial explanation level and self-relevance between affective stimuli alone can lead to the paradox. As shown by simulations using HED by its author, HED explains the paradox successfully only when an intense emotion is involved whereas HED-ID demonstrates that the paradox can occur across the full range of emotion intensities, which is more in line with empirical observations.

One of the major departures of HED-ID from both the models is the introduction and definition of the explanation level. The understanding one has of an event, as defined by the AREA model, is split into the separate parameters of explanation level and explanatory ease. Whereas explanatory ease represents the ease with which one could explain the event, the level of anticipation of the event is the dominant determinant of the explanation level. For an unanticipated event, the explanation level at the time of event occurrence is always zero irrespective of the event's self-relevance and how hard or easy it is to explain. If the event is anticipated, the explanation level at the time of its occurrence depends on how early it was anticipated, with how much certainty it is anticipated and how easy it is to explain the stimulus. We believe, this segregation has helped us to provide a more accurate explanation to empirical observations and make more specific empirically testable predictions.

For instance, in all the empirical instances of Scenario 1, the AREA model would just predict a stronger and longer emotional response for the stimulus with less certainty or coherence. HED-ID is more accurate with its predictions. It predicts that the emotion elicited by both the stimuli would be similar initially but that the emotion resulting from the less coherent or uncertain stimulus would last longer which is more consistent with the actual observations. To take another example, for the study by Rasinski et al. [15], which is an instance of Scenario 3, AREA would predict a higher initial emotional impact for the more affected group due to the event being of greater relevance to them, consistent with the observations of the authors. But, contradictory to what they observed, it would also predict a longer duration of emotional impact for the more affected group because it would not be easier for them to explain the event compared to the less affected group. This is also the case with other similar studies demonstrating the region-β paradox such as those in [14]. On the other hand, the predictions of HED-ID are fully consistent with the observations in all these studies. Whereas the AREA model provided a general framework to explain affective adaptation and the resulting emotional response, HED-ID is able to offer more targeted predictions and explain specific observations. Being quantitative in nature, in addition to emotion duration and intensity, HED-ID can predict how experienced emotion changes over time due to affective adaptation. Further, the model can be used in artificial systems that require affectively intelligent agents.

4.3 Region-β paradox

The only direct demonstration of the region-β paradox in the laboratory till date has been study 2 of Gilbert et al. [14]. That study's limitations, such as the use of small sample sizes and different dependent variables for the 2a and 2b parts, coupled with the lack of other relevant research may question the actual existence of the phenomenon in the real world. However, since emotional responses of the same intensity can adapt at different rates, as empirical data have shown, one can directly infer that it is possible for a larger emotional response to adapt completely before a smaller one if it adapts sufficiently faster. Further, HED-ID computationally shows that the paradox can easily happen under different circumstances and for different reasons. Thus one may conclude that the paradox, though fascinating, is not surprising and is a likely regular occurrence in everyday life. Indeed, a real-life example of the paradox can be seen in Rasinski et al. [15] as discussed earlier.

Gilbert et al. [14] had proposed that the paradox would arise only when an intense emotion is involved. As pointed out in the Introduction, empirical data is inconsistent with this proposal. Simulations with HED-ID suggest that the paradox can happen within the full range of emotion intensities. HED-ID also proposes that the paradox is not a function of intensities alone as suggested by Gilbert et al. [14] and Steephen [27]. It is more of a function of the relative levels of understanding and self-relevance as well as the ease of explaining the stimulus. Our simulations further showed that depending on conditions, the *cross-over* associated with the paradox can happen either early or late in the adaptation process. However, the reader needs to be cautioned that even under the conditions for which HED-ID predicts the paradox, it would not occur if the emotion from the weaker stimulus is so small that it dies down before getting a chance to cross over. We hope this paper motivates further research into this interesting phenomenon.

4.4 Applications

Development of computational models of emotion serves at least two purposes. One is to help understand the emotion system and emotion processing better. The other is to apply in artificial systems. For instance, an analogue of emotion is sometimes required during decision making in an artificial organism [34], [35], [36], [37]. When the goal is the former, the focus would be to make the system as close to reality as possible whereas believability may be more important in the latter. While the primary purpose of HED-ID is to gain more insights into emotion psychology, we believe it can be usefully employed in developing affectively intelligent agents that can be used in areas such as virtual reality

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TAFFC.2018.2848656, IEEE Transactions on Affective Computing

training, games, human-computer interactions, and robotics. The simplicity and computational cost effectiveness of HED-ID can be taken advantage of in such applications. See Steephen [27] for examples of how models such as HED-ID can be used beneficially in different affective computing environments.

4.4 Limitations and future directions

Although retrospective comparisons show that the model predictions are consistent with existing empirical data, a stronger validation of HED-ID would come only when controlled experiments designed specifically to test it are conducted. One may wonder as to whether it is possible to create the different scenarios empirically by controlling the values of the relevant psychological variables of the model. Though quantities such as self-relevance, explanation level and explanatory ease cannot always be accurately compared between stimuli, one could design stimuli where these can be clearly made to differ. For instance, for most people, winning $1000 is more self-relevant than winning $100. Initial explanation level can be manipulated by providing advance information of the stimulus to half the subjects. Experiments similar to a more controlled variant of our empirical study could be used to manipulate explanatory ease. For instance, the same clip can be shown to all the subjects, but half of them may be informed that it is fiction whereas the others told otherwise.

A limitation of HED-ID is that it assumes the amount of cognitive resources allotted to affective adaptation to be a function of emotion intensity alone. While one may get increasingly motivated to enhance the adaptation effort exerted with increasing emotion intensity, and this may be what happens most of the time, it is important to note that there could be additional factors that influence this effort. For instance, Larson & Sbarra [38] reported that just participating in a research assessing coping, that involved only measurements and no interventions, promoted emotional recovery. Here, participation in the research is unlikely to have changed the self-relevance or explanatory ease. However, being assessed would make one allocate more attentional resources to the event, thereby increasing the cognitive effort employed for adaptation. Secondly, an increased cognitive effort need not always translate into faster adaptation as assumed by HED-ID. Certain coping strategies such as rumination that have been reported to increase in amount and duration with affect intensity [8], [16], if sustained for too long may end up prolonging the emotional experience. Third, anticipating a future affective stimulus starts the process of explanation and adaptation even before its actual occurrence. HED-ID does not predict the variations in explanation level and emotional response during this anticipation phase. It should be noted that empirical as well as theoretical work on anticipatory adaptation is virtually non-existent and is an area that deserves greater attention. Fourth, HED-ID does not model the net emotional experience of overlapping affective stimuli. For instance, if the conditions are such that two sad events occurring at different times in an individual's life lead to the region-β paradox, would the paradox happen if they were to occur simultaneously, that is, would the individual adapt faster to the sadder event? Fifth, HED-ID, in its current form, cannot be used with multidimensional emotion models (e.g. [39], [40]). A HED-ID variant incorporating emotion's underlying dimensions, in which the time course of each dimension is tracked separately, could be a useful addition to literature. Sixth, we assumed that linear relationships between the HED-ID parameters would be parsimonious on the basis of available data and literature. However, in reality, it is possible that these relationships are more complex. HED-ID would need appropriate revisions once the nature of these relationships become clearer from future empirical studies. Finally, how HED-ID relates to the psychopathologies and psychotherapeutic approaches related to emotion was not examined and is a potential area for further exploration.

5 Conclusion

We explored the relationship between emotion intensity and its duration using empirical and computational approaches and also addressed the shortcomings of past affective adaptation models. We proposed the HED-ID model, described by three algebraic equations and three independent variables, by extending and quantifying the concepts underlying the AREA and HED models. HED-ID addressed the intensity-duration problem of emotion by proposing the factors that determine emotion intensity and duration and how they individually influence these aspects. HED-ID predicted the conditions under which the correlation between intensity and duration is positive, negative and absent and helped explain the conflicting observations on the intensity-duration relationship. We also showed that the model is consistent with past empirical observations. Our empirical work, supported the modeling effort by successfully validating the hypothesis that emotion intensity is not the sole predictor of its duration, thereby strengthening the notion that these two aspects are determined differently by the stimulus. We hope that future research would address the limitations of HED-ID and facilitate further understanding of affective adaptation and emotion dynamics.

Acknowledgment

This study was supported by Grant SR/CSI/PDF-21/2013 from the Department of Science and Technology, New Delhi. We thank the Kerala University Library staff, especially Mr. Hamsa, for their help in writing this paper. Special thanks to Timothy Wilson for his comments that helped improve this paper. J. E. Steephen is the corresponding author.

References

[1] J. J. Gross and R. W. Levenson, "Emotion elicitation using films," *Cogn. Emot.*, vol. 9, no. 1, pp. 87–108, Jan. 1995.

[2] C. A. Smith and R. S. Lazarus, "Emotion and adaptation," in *Handbook of Personality: Theory and Research*, L. A. Pervin, Ed. New York: Guilford Press, 1990, pp. 609–637.

[3] P. C. Ellsworth and K. R. Scherer, "Appraisal processes in emotion," in *Handbook of Affective Sciences*, R. J. Davidson and K. R. Scherer, Eds. New York: Oxford University Press, 2003, pp. 572–595.

[4] K. R. Scherer, "The nature and study of appraisal: A review of the issues," *Apprais. Process. Emot. Theory, methods, Res.*, pp. 369–391, 2001.

[5] J. Sonnemans and N. H. Frijda, "The structure of subjective emotional intensity," *Cogn. Emot.*, vol. 8, no. 4, pp. 329–350, 1994.

[6] T. D. Wilson and D. T. Gilbert, "Explaining away: A model of affective adaptation," *Perspect. Psychol. Sci.*, vol. 3, no. 5, pp. 370–386, 2008.

[7] N. H. Frijda, B. Mesquita, J. Sonnemans, and S. Van Goozen, "The duration of affective phenomena or emotions, sentiments and passions," in *International Review of Studies on Emotion*, vol. 1, K. T. Strongman, Ed. Chichester, England: Wiley, 1991, pp. 187–225.

[8] N. H. Frijda, *Laws of Emotion*. Mahwah, NJ: Erlbaum, 2007.

[9] M. Luhmann, W. Hofmann, M. Eid, and R. E. Lucas, "Subjective well-being and adaptation to life events: a meta-analysis.," *J. Pers. Soc.*

Psychol., vol. 102, no. 3, pp. 592–615, Mar. 2012.
[10] S. Lyubomirsky, “Hedonic adaptation to positive and negative experiences,” in *The Oxford Handbook of Stress, Health, and Coping*, S. Folkman, Ed. New York: Oxford University Press, 2010, pp. 200–224.
[11] S. Frederick and G. F. Loewenstein, “Hedonic adaptation,” in *Well-Being: The Foundations of Hedonic Psychology*, D. Kahneman, E. Diener, and N. Schwarz, Eds. New York: Russell Sage Foundation, 2003, pp. 302–329.
[12] M. A. Strahilevitz and G. F. Loewenstein, “The effect of ownership history on the valuation of objects,” *J. Consum. Res.*, vol. 25, no. 3, pp. 276–289, Dec. 1998.
[13] D. T. Gilbert, E. Driver-Linn, and T. D. Wilson, “The trouble with Vronsky: Impact bias in the forecasting of future affective states,” in *The Wisdom in Feeling: Psychological Processes in Emotional Intelligence*, L. F. Barrett and P. Salovey, Eds. New York: Guilford Press, 2002, pp. 114–143.
[14] D. T. Gilbert, M. D. Lieberman, C. K. Morewedge, and T. D. Wilson, “The peculiar longevity of things not so bad,” *Psychol. Sci.*, vol. 15, no. 1, pp. 14–19, 2004.
[15] K. A. Rasinski, J. Berktold, T. W. Smith, and B. L. Albertson, “America recovers: A follow-up to a national study of public response to the September 11th terrorist attacks,” Chicago, 2002.
[16] B. Rimé, P. Philippot, S. Boca, and B. Mesquita, “Long-lasting cognitive and social consequences of emotion: Social sharing and rumination,” in *European Review of Social Psychology*, vol. 3, W. Stroebe and M. Hewstone, Eds. New York: Wiley, 1992, pp. 225–258.
[17] E. Aronson and J. Mills, “The effect of severity of initiation on liking for a group,” *J. Abnorm. Soc. Psychol.*, vol. 59, no. 2, pp. 177–181, 1959.
[18] P. G. Zimbardo, A. R. Cohen, M. Weisenberg, L. Dworkin, and I. Firestone, “Control of pain motivation by cognitive dissonance,” *Science (80-.).*, vol. 151, no. 3707, pp. 217–219, Jan. 1966.
[19] H. B. Gerard and G. C. Mathewson, “The effects of severity of initiation on liking for a group: A replication,” *J. Exp. Soc. Psychol.*, vol. 2, no. 3, pp. 278–287, 1966.
[20] S. E. Brandon and A. P. Silke, “Near and long-term psychological effects of exposure to terrorist attacks,” in *Psychology of Terrorism*, B. Bongar, L. M. Brown, L. E. Beutler, J. N. Breckenridge, and P. G. Zimbardo, Eds. New York: Oxford University Press, 2007, pp. 175–193.
[21] P. Smith, S. Perrin, W. Yule, and S. Rabe-Hesketh, “War exposure and maternal reactions in the psychological adjustment of children from Bosnia-Hercegovina.,” *J. Child Psychol. Psychiatry.*, vol. 42, no. 3, pp. 395–404, Mar. 2001.
[22] A. A. M. Thabet and P. Vostanis, “Post-traumatic stress reactions in children of war,” *J. child Psychol. Psychiatry*, vol. 40, no. 3, pp. 385–391, 1999.
[23] H. Helson, *Adaptation-Level Theory*. New York: Harper & Row, 1964.
[24] V. Quinsey, “Some applications of adaptation-level theory to aversive behavior,” *Psychol. Bull.*, vol. 73, no. 6, pp. 441–450, 1970.
[25] J. L. Kurtz, T. D. Wilson, and D. T. Gilbert, “Quantity versus uncertainty: When winning one prize is better than winning two,” *J. Exp. Soc. Psychol.*, vol. 43, no. 6, pp. 979–985, 2007.
[26] T. D. Wilson, D. B. Centerbar, D. A. Kermer, and D. T. Gilbert, “The pleasures of uncertainty: Prolonging positive moods in ways people do not anticipate.,” *J. Pers. Soc. Psychol.*, vol. 88, no. 1, pp. 5–21, Jan. 2005.
[27] J. E. Steephen, “HED: A computational model of affective adaptation and emotion dynamics,” *IEEE Trans. Affect. Comput.*, vol. 4, pp. 197–210, 2013.
[28] D. P. Lovell and F. D. Zeffirelli, *The champ*. USA: MGM, 1979.
[29] J. Zaritsky, *The suicide tourist*. Canada: CTV Television Network, Point Grey Pictures Inc., 2007.
[30] J. W. Peirce, “PsychoPy - Pyschophysics software in Python,” *J. Neurosci. Methods*, vol. 162, no. 1–2, pp. 8–13, 2007.
[31] D. T. Gilbert, E. C. Pinel, T. D. Wilson, S. J. Blumberg, and T. P. Wheatley, “Immune neglect: A source of durability bias in affective forecasting.,” *J. Pers. Soc. Psychol.*, vol. 75, no. 3, pp. 617–38, Sep. 1998.
[32] P. Brickman and D. T. Campbell, “Hedonic relativism and planning the good society,” in *Adaptation-level Theory: A Symposium*, M. H. Appley, Ed. New York: Academic Press, 1971, pp. 287–302.
[33] S. Wiggins, P. Whyte, M. Huggins, S. Adam, J. Theilmann, M. Bloch, S. Sheps, and M. Hayden, “The psychological consequences of predictive testing for Huntington’s disease,” *N. Engl. J. Med.*, vol. 327, no. 20, pp. 1401–1405, 1992.
[34] R. S. Bapi, “Exploring the role of emotion in the design of autonomous systems,” in *Proceedings of the Helsinki Neural Network Group Workshop HELNET*, 1996, vol. 95, pp. 4–7.
[35] J. Broekens, D. Degroot, and W. A. Kosters, “Formal models of appraisal: Theory, specification, and computational model,” *Cogn. Syst. Res.*, vol. 9, no. 3, pp. 173–197, 2008.
[36] E. Hogewoning, J. Broekens, J. Eggermont, and E. G. P. Bovenkamp, “Strategies for affect-controlled action-selection in Soar-RL,” *Lect. Notes Comput. Sci.*, vol. 4528, pp. 501–510, 2007.
[37] J. Broekens, W. A. Kosters, and F. J. Verbeek, “Affect, anticipation, and adaptation: Affect-controlled selection of anticipatory simulation in artificial adaptive agents,” *Adapt. Behav.*, vol. 15, no. 4, pp. 397–422, 2007.
[38] G. M. Larson and D. A. Sbarra, “Participating in research on romantic breakups promotes emotional recovery via changes in self-concept clarity,” *Soc. Psychol. Personal. Sci.*, vol. 6, no. 4, pp. 399–406, 2015.
[39] J. R. J. Fontaine, K. R. Scherer, E. B. Roesch, and P. C. Ellsworth, “The World of Emotions Is Not Two-Dimensional,” *Psychol. Sci.*, vol. 18, no. 12, pp. 1050–1057, 2015.
[40] J. Russell, “A circumplex model of affect,” *J. Pers. Soc. Psychol.*, vol. 39, no. 6, pp. 1161–1178, 1980.

John E. Steephen received the BTech degree in applied electronics and instrumentation engineering in 1996 from College of Engineering Trivandrum. He received the MTech degree in biomedical engineering in 2004 and the PhD degree in computational neuroscience in 2011 from IIT Bombay. He has been a DST-CSI postdoctoral fellow at the Centre for Neural and Cognitive Sciences, University of Hyderabad and KCIS, IIIT Hyderabad. His current work involves empirical and computational explorations of emotion perception and experience.

Siva C. Obbineni has an MA (Communication) degree from University of Hyderabad and has been an MPhil student at the Centre for Neural and Cognitive Sciences, University of Hyderabad.

Sneha Kummetha is a BTech-MS dual degree student at Cognitive Science Lab, KCIS, IIIT Hyderabad.

Raju S. Bapi received the BE (Electrical Engineering) degree in 1983 from Osmania University. He received the MS (Biomedical Engineering) and PhD (Mathematical Sciences Computer Science) degrees from the University of Texas at Arlington. He worked at BHEL; University of Plymouth; and ATR Research Labs, Kyoto before joining University of Hyderabad. He has been a Visiting Professor at KCIS, IIIT Hyderabad. His research activities involve the practical applications of various neural network and machine learning techniques, investigation of biological neural architectures and cognitive modelling.