# A database of news videos for investigating the dynamics of emotion and memory

Rosalie Samide ¹, Rose A. Cooper ¹, & Maureen Ritchey \*¹

¹ Department of Psychology, Boston College

\* Correspondence concerning this article should be addressed to Maureen Ritchey, Boston College Department of Psychology, 140 Commonwealth Ave, Chestnut Hill, MA 02467. Email: maureen.ritchey@bc.edu. Phone: +1 617 552 4779. Fax: +1 617 552 0523.

# Abstract

Emotional experiences are known to be both perceived and remembered differently from non-emotional experiences, often leading to heightened encoding of salient visual details and subjectively vivid recollection. The vast majority of previous studies have used static images to investigate how emotional event content modulates cognition, yet natural events unfold over time. Therefore, little is known about how emotion dynamically modulates continuous experience. Here, we report a norming study wherein we develop a new stimulus set of 126 emotionally negative, positive, and neutral videos depicting real-life news events. Participants continuously rated the valence of each video during its presentation and judged the overall emotional intensity and valence at the end of each video. In a subsequent memory test, participants reported how vividly they could recall the video details and estimated each video’s duration. We report data on the affective qualities and subjective memorability of each video. The results replicate the well established effect that emotional experiences are more vividly remembered than non-emotional experiences. Importantly, this novel stimulus set will facilitate research into the temporal dynamics of emotional processing and memory.

# INTRODUCTION

Everyday experiences contain a wealth of information, from perceptual details to thoughts and emotions. Cognitive psychology and neuroscience research aims to uncover the basic cognitive and neural processes supporting perception and memory of complex events. Such investigations, however, are faced with a tension between experimental control and ecological validity: in order to isolate a single process, it is often necessary to use very simple stimuli. Yet this approach misses other defining aspects of experience, such as the integration of multimodal streams of information and dynamic changes in these streams-- and our reactions to them-- over time. Therefore, basic research on cognitive and neural processes can benefit from complementary approaches that incorporate dynamic stimuli to improve generalizability to real-life events.

Natural experiences often include affective elements or states that vary over time. Past research on emotion processing has largely focused on one of two extremes: transient changes in affect elicited by static stimuli, such as words (e.g., Affective Norms for English Words (ANEW); Bradley & Lang, 2017) or visual scenes, including the well known International Affective Picture System (IAPS; Lang et al., 2008) and the Nencki Affective Picture System (NAPS; Marchewka et al. 2014), or sustained changes in affect elicited by mood inductions or stressors (e.g., Richell & Anderson, 2004; Stroud et al., 2000). Studies taking the former approach have shown that in addition to increasing subjective feelings of memory vividness (for review see Phelps & Sharot, 2008), emotion strongly influences the way information is perceived (for review see Zadra & Clore, 2011) and the types of details that are recalled (for review see Kensinger, 2009; Yonelinas & Ritchey, 2015). For instance, testing memory for emotional images, such as from the IAPS database, shows that negative items tend to be remembered in more detail (Kensinger et al., 2007), and the presence of emotion increases activity in visual cortex (Lang et al., 1998; Kark & Kensinger, 2015). In scenes that combine emotional items with neutral background information, emotion tends to enhance memory for central emotional information while impairing memory for peripheral scene details (Kensinger et al., 2007; Bisby & Burgess, 2013), suggesting that even in static images, the effects of emotion can be quite specific to particular image features.

Although scene stimuli are visually complex, they lack the temporal dynamics of real-life experiences. As such, it remains unclear how the time course of emotional experience affects the processing of complex events. Outside of the emotion domain, naturalistic video stimuli (e.g., clips from movies or television shows) have been increasingly used to investigate the dynamic nature of cognitive and neural processes during perception and subsequent episodic memory retrieval. For example, video stimuli have been used to demonstrate similarities in episodic representation between perception and retrieval (Bird et al., 2015; Chen et al., 2017; Oedekoven et al., 2017). Video stimuli also have provided novel insight into how the brain codes for event boundaries and represents timescales (Baldassano et al., 2017). These findings have significantly furthered our understanding of how the brain represents and reconstructs complex events and could not have been ascertained using static images. There has also been a push within social psychology and neuroscience to use naturalistic stimuli (for review see Risko et al., 2012) to study, for example, how interpersonal dynamics develop (Curio et al., 2010; Klin et al., 2002) and influence our perception and interpretation of the world (Putman et al., 2006). Thus, research across the broad fields of psychology and neuroscience are in need of ecologically valid, dynamic stimuli.

While research using temporally dynamic stimuli has provided novel insight into cognitive and neural processes of episodic memory, in these studies emotional content has not been directly and systematically measured or manipulated. For example, though the Sherlock database includes valence and arousal ratings collected from several experimenters (Chen et al., 2017), it does not include ratings from a naive sample of subjects over whom ratings could be averaged to increase generalizability to the population. Within the emotion literature, there are a limited number of video stimulus databases that have been developed to date. The Database for Emotion Analysis Using Physiological Signals (DEAP, Koelstra et al., 2012) is comprised of minute-long clips from music videos that were rated for valence, arousal, familiarity, dominance, and like versus dislike while electroencephalographic (EEG) data were collected. DEAP is thus optimal for other researchers to use as a comparison against other methods to measure affective states, but music videos are not an ideal proxy for complex, real-life experiences because they contain music rather than dialogue and often lack a realistic storyline. The MAHNOB-HCI database (Soleymani et al., 2012) includes short film clips taken from well-known emotion-inducing movies such as The Shining and Gangs of New York, also with accompanying EEG recordings. Similarly, the DECAF database (Abadi et al., 2015) includes data on behavioral, magnetoencephalographic (MEG), and other physiological responses to stimuli from both DEAP and MAHNOB-HCI. DECAF also contains dynamic annotations indicating the intended evoked emotion made by experts who had viewed the film clips numerous times. Another multimodal stimulus set, the FilmStim database (Shaefer et al., 2010), is comprised of 1-7 minute-long film segments from movies that were shown to elicit the basic emotions of anger, sadness, fear, and disgust (Ekman, 1984), as well as amusement, tenderness, and a neutral state. These existing stimuli datasets have some limitations, which have motivated the current project. First, film clips taken from movies, television shows, or music videos may be processed differently than autobiographical events because participants are aware that the events depicted are fictional (Abraham et al., 2008). Additionally, though other video databases have measured emotion, this has typically been rated over short segments, thus averaging over a subsection of the full-length stimulus. A dynamic or continuous measure of participants’ emotional experiences while watching stimuli would be useful for future experiments.

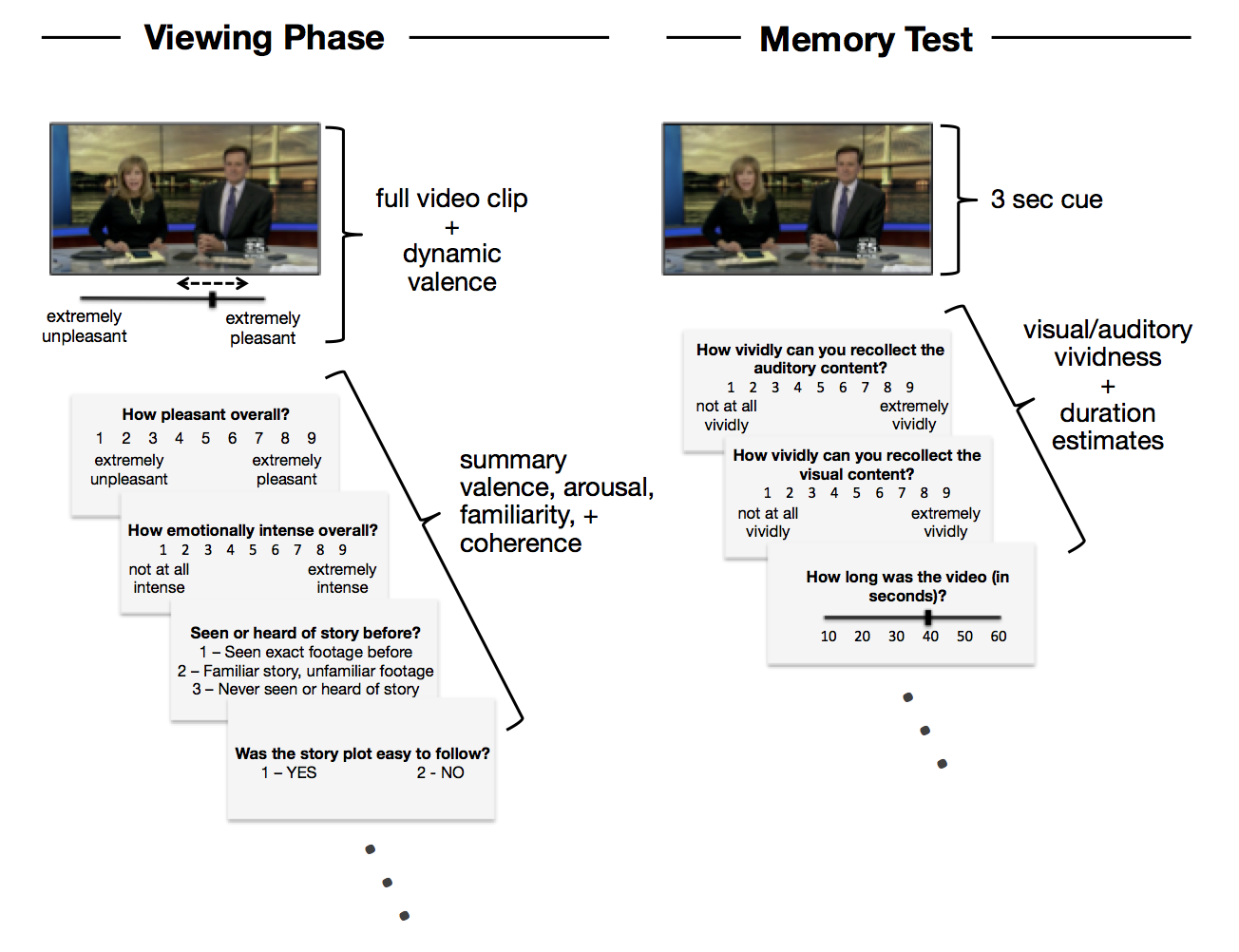
The goal of the current study was to develop a new set of natural video stimuli, selected and rated for their affective content, that will enable future research on the dynamics of emotion processing and memory. To improve the ecological validity of this set, we focused on clips from real-life news telecasts covering stories ranging from terrorist attacks and deadly car accidents to military homecomings and rescued animals. To account for the dynamic nature of the stimuli, participants continuously rated the valence of the video throughout its presentation and then judged how emotionally intense (arousal judgement) and pleasant (summary valence judgement) each video was overall. Furthermore, a basic memory test allowed us to collect subjective memorability ratings for each video, which we hope will be useful for future studies aiming to directly contrast ‘memorable’ and ‘forgettable’ experiences (cf. Bainbridge, 2017). Thus, this stimulus set will allow researchers to assess memory for natural, real-life experiences and how accompanying cognitive and neural processes are dynamically modulated by emotion.

# METHODS

## Participants

A total of 100 participants (69 females, 31 males) took part in the current experiment. All participants were between the ages of 18 and 22 (mean = 19.01, SD = 0.85), had normal or corrected-to-normal vision, and had no current diagnoses or history of psychological or neurological disorders. Informed consent was obtained from all participants. All procedures were approved by the Boston College Institutional Review Board, and participants received course credit for their time.

## Materials

The videos were collated from a television news archive found at<https://archive.org/details/tv>. Videos were downloaded based on searches for keywords or phrases in the transcript. To get a range of negative videos, keywords included words such as ‘disaster’, ‘murder’, ‘poverty’, ‘victim’, and ‘tragic’. Neutral videos including words such as ‘weather’, ‘traffic’, ‘school’, ‘construction’, and ‘business’, and positive videos were found with 

*Figure 1.* Task design. During the Viewing Phase, participants watched news video clips while continuously rating pleasantness. After each clip, they then rated summary pleasantness (valence), emotionally intensity, familiarity with the story, and story coherence. For the Memory Test, which was interleaved with the Viewing Phase in blocks, participants were cued to retrieve each video with a 3 second clip and then judged the vividness of their memory for the auditory content, visual content, and estimated the full video clip length.

words including ‘happy’, ‘celebration’, ‘heartwarming’, ‘surprise’, and ‘uplifting’. These videos were then manually filtered to remove any with low quality resolution, any containing an event similar to other videos, and any reporting highly familiar, international news stories. Videos containing only a verbal description of the event were also excluded; thus, the remaining videos all contained some visual footage of the event. This process resulted in a total of 144 videos that were selected to be used in the norming experiment, 48 of which came from negative keywords, 48 from neutral keywords, and 48 from positive keywords, based on keyword valence judged by the experimenters. All videos were trimmed to be between 20 and 52 seconds in duration and to remove any footage at the beginning or end that did not pertain to the central news story. The mean duration was 42.15 seconds (SD = 7.70). All videos were 640 pixels in width and between 360 and 480 pixels in height.

## Procedure

After giving informed consent, participants completed the viewing phase of the experiment. The 144 videos were divided into two lists of 72; each list contained 24 videos of each predefined valence category. Each participant viewed the videos from one of the two lists to limit the length of the experimental session, which lasted approximately 1 hour 30 minutes. Lists were assigned to participants in an alternating fashion. Therefore, 50 participants rated each of the 144 videos. The order of the videos was randomized per participant and the session was divided into 4 study-test blocks, each containing 18 videos, to allow for rest breaks. For each study trial, participants first watched the video while adjusting a continuous slider to indicate how pleasant the video was to them at that moment in time on a scale of ‘extremely unpleasant’ (coded as 1) to ‘extremely pleasant’’ (coded as 9). This ‘dynamic valence*’* slider started at a random position on the scale at the beginning of each trial, and participants were asked to adjust the slider as quickly as possible to reflect their impression of the video and to keep adjusting it as the perceived pleasantness changed. At the end of each video, participants were asked an additional 4 questions: overall valence (‘summary valence’) on a scale of 1 (extremely negative) to 9 (extremely positive), overall arousal (‘intensity’) on a scale of 1 (not at all intense) to 9 (extremely intense), the familiarity of the video with the options of ‘I have seen this exact news footage before’, ‘I am familiar with the news story but I have not seen this footage’, and ‘I have not seen or heard of this news story before’, and finally if the story depicted in the video was coherent or ‘easy to follow’ with ‘yes’ or ‘no’ response options.

After watching all 18 videos in the study phase of a block, participants completed a memory test. On each test trial, participants were first shown the first 3 seconds of the video as a retrieval cue. After this clip, they were asked to rate how vividly they could recollect the auditory details about the video and the visual details of the video, both using the scales 1 (not at all vividly) to 9 (extremely vividly). Participants were instructed to try to remember as much of the video content as possible before responding. They were then shown a scale from 10 seconds to 60 seconds with tick marks at 10 second intervals and were asked to estimate the total duration of the video by moving the slider along the scale and pressing the spacebar to confirm their response. The slider appeared a random location along this duration scale on every trial. All responses were self-paced.

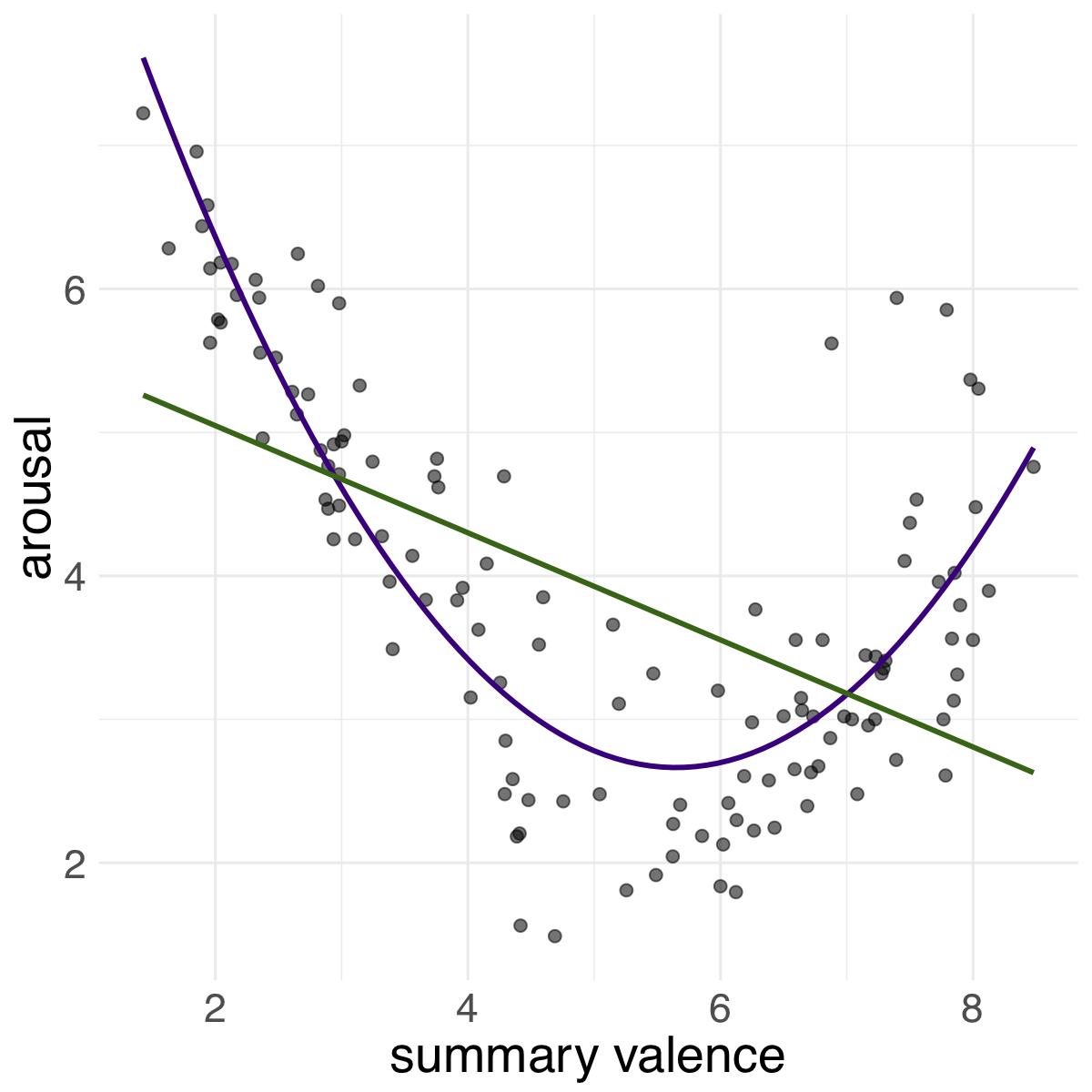
# RESULTS

Videos were excluded prior to data analysis if more than 25% of participants rated the video as incoherent (16 videos). Two additional videos were removed from data analyses as a result of experimenter error: one because its duration was actually significantly longer than 52 seconds, and one because it showed almost identical news footage to a different video. The remaining sample of 126 videos (mean duration = 42.43 seconds, SD = 7.44 seconds) are included in all analyses below. All data for the videos can be found in Supplemental 1 & 2 and at [http://www.thememolab.org/paper-videonorming](http://www.thememolab.org/paper-videonorming/), including the mean of all measures collected from participants - dynamic and summary ratings of emotional valence and arousal, ratings of video familiarity, and mean memory vividness and duration estimates. We additionally include information coded by the researchers, including semantic categories, number of people present, location, year, as well as short and long unique video descriptions. Although we refrain from directly distributing the videos due to copyright issues, we have provided the URLs for each video and example code that other researchers can use to download the same videos from the TV news archive for use in their research. All analyses were performed using R Version 3.5.0 with RStudio Version 1.1.456 (R Core Team, 2012).

## Summary Valence and Arousal Ratings

Summary valence and arousal ratings collected at the end of each video were averaged across participants, and the distribution of these mean ratings can be seen in Figure 3a. We first analyzed the relationship between these ratings across the videos. Mean valence and mean arousal displayed a stereotypical asymmetric v-shaped relation (for review see Kuppens et al., 2013). To quantify this relationship, we used R’s lme4 package (Bates, Maechler & Bolker, 2012) to perform a linear mixed effects analysis of the relationship between arousal and valence. As fixed effects, we entered valence and its square into the model. We included subject and video as random effects by allowing the intercept to vary. Arousal was the dependent variable. P-values were obtained by likelihood ratio tests of the full model with the effect in question against the model without the effect in question. Valence was related to arousal both linearly ((1) = 444.3, *p* < .001), 𝛽 = 1.10 *±* 0.05 (standard error) and quadratically ((1) = 437.57, *p* < .001), 𝛽 = 0.10 *±* 0.005, such that videos on the more negative and positive ends of the valence spectrum were rated as more intense than neutral videos, but negative videos tended to be the most emotionally intense.

Next we analyzed the stability of summary valence and arousal across subjects. To this end, we iteratively correlated the vector of each participant’s summary valence ratings for each video with the summary valence ratings of every other participant who had watched the same videos. The mean of these correlations (r = 0.74) reflects a high level of consistency of perceived valence across subjects. Repeating this process with arousal ratings revealed that emotional intensity was also stable across subjects, r = 0.40, although to a lesser degree than valence. To determine which videos were the most variable in their ratings across subjects, we calculated the standard deviation of summary valence (mean sd = 1.35) and arousal (mean sd = 2.05) across subjects for each video. These data are reported in Supplemental 1. Notably, the stability of the valence and arousal measures did not vary according to their mean values, indicated by multiple regression analyses revealing no significant linear or quadratic relationship between mean summary valence and valence ambiguity (*p*s > .125), or between mean summary arousal and arousal ambiguity (*p*s > .145).

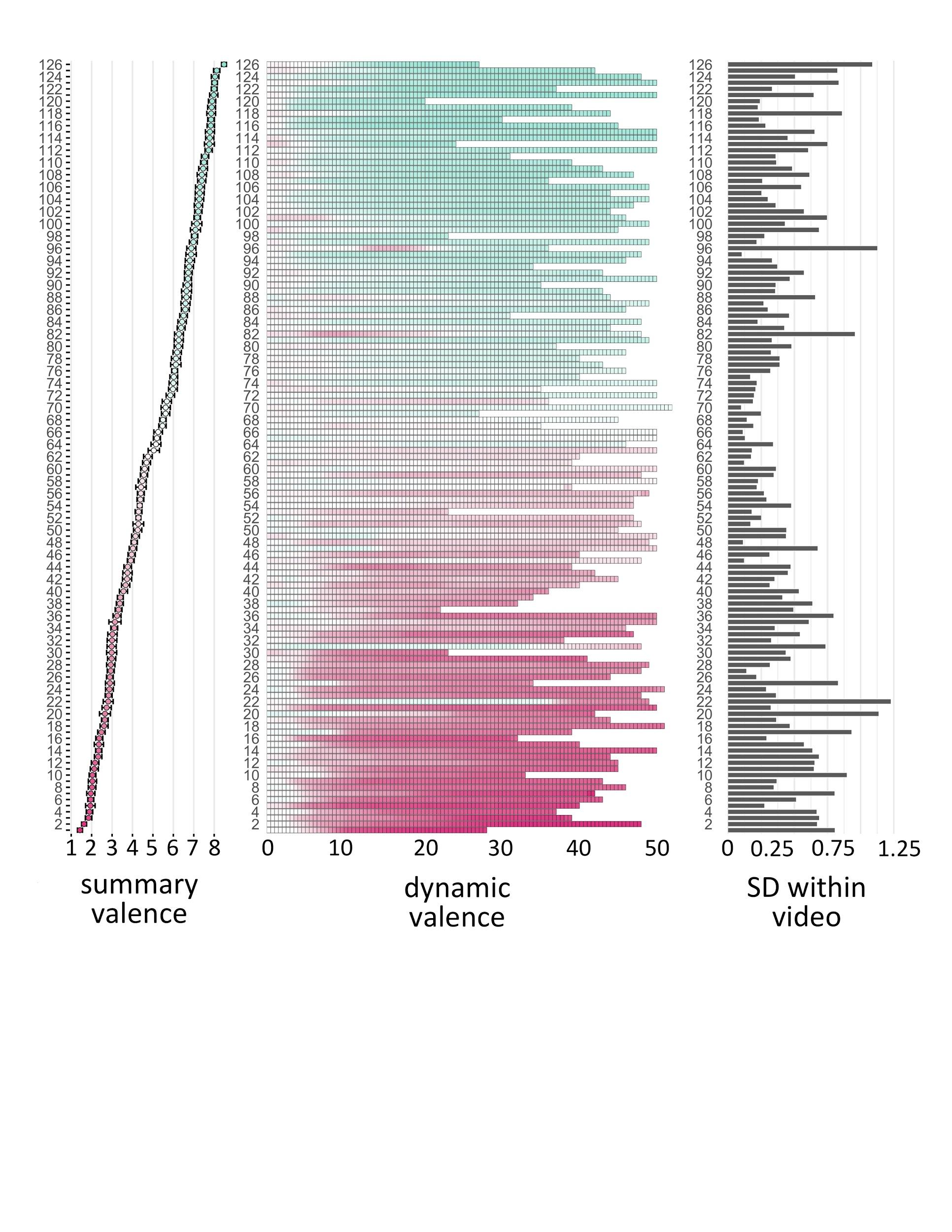


*Figure 2.* Each video plotted by mean ratings of arousal (1 = least intense, 9 = most intense) and valence (1 = most negative, 9 = most positive). Green line represents the linear relation between measures; purple line represents the quadratic relationship between measures.

## Dynamic Pleasantness Ratings

We next investigated the properties of the dynamic valence ratings obtained for each video. In order to examine the time course of these ratings, the mean rating across participants was calculated for each video at every 500ms time point (Figure 3b). To examine how much the valence of each video varied over time, we then calculated the standard deviation of those time point means for each video (Figure 3c). The first 5 seconds of dynamic valence ratings were excluded from all calculations since the valence slider appeared in a random location at the start of each video and participants took a few seconds to move the slider based on their first impression of the video content. Standard deviation of dynamic valence was strongly correlated with video arousal ratings, r(124) = .56, *p* < .001, illustrating that negative and positive videos were associated with more variability in continuous valence ratings over the video time course. This finding may reflect the gradual accumulation of information that leads participants to shift from a relatively neutral impression of the video to a more clearly positive or negative impression for the more emotionally arousing videos. Standard deviation of dynamic valence was not significantly correlated with summary valence (r < 0.20, *p* > .20).

To analyze the correspondence between dynamic valence ratings and summary valence ratings, we first correlated the mean of dynamic valence ratings with mean summary valence ratings for each video. This revealed a strong relationship, r(124) = 0.96, *p* < .001. To examine whether this correlation was driven by the beginning, middle, or end of each video, we divided each video’s mean dynamic time course into thirds, and correlated the mean dynamic valence of each third with the mean summary valence. The mean dynamic valence from the beginning third was strongly correlated with summary valence, r(124) = 0.89, *p* < .001. This relationship increased for mean dynamic valence within the middle third, r(124) = 0.95, *p* < .001, and the end third, r(124) = 0.99, *p* < .001, which had the highest correlation between dynamic and summary valence ratings. The peak of dynamic valence ratings was also calculated across subjects for each video by centering the valence scale about 0 and taking the maximum absolute value at any time point for each video. Similarly, peak dynamic valence was highly correlated with mean summary valence ratings (r(124) = 0.97, *p* < .001).

**A B C**

*Figure 3.* Valence ratings. Each row represents data for a single video in order of (A) ascending mean summary valence. Participants rated summary valence on a scale from 1-9 (mean ratings ranged from 1.43 to 8.48). (B) Mean dynamic valence rating across subjects at each 500ms timepoint. Pink represents more negative ratings, turquoise represents more positive ratings. (C) Standard deviation of mean dynamic valence ratings over time, representing how much valence was perceived to fluctuate within each video.

## Memorability

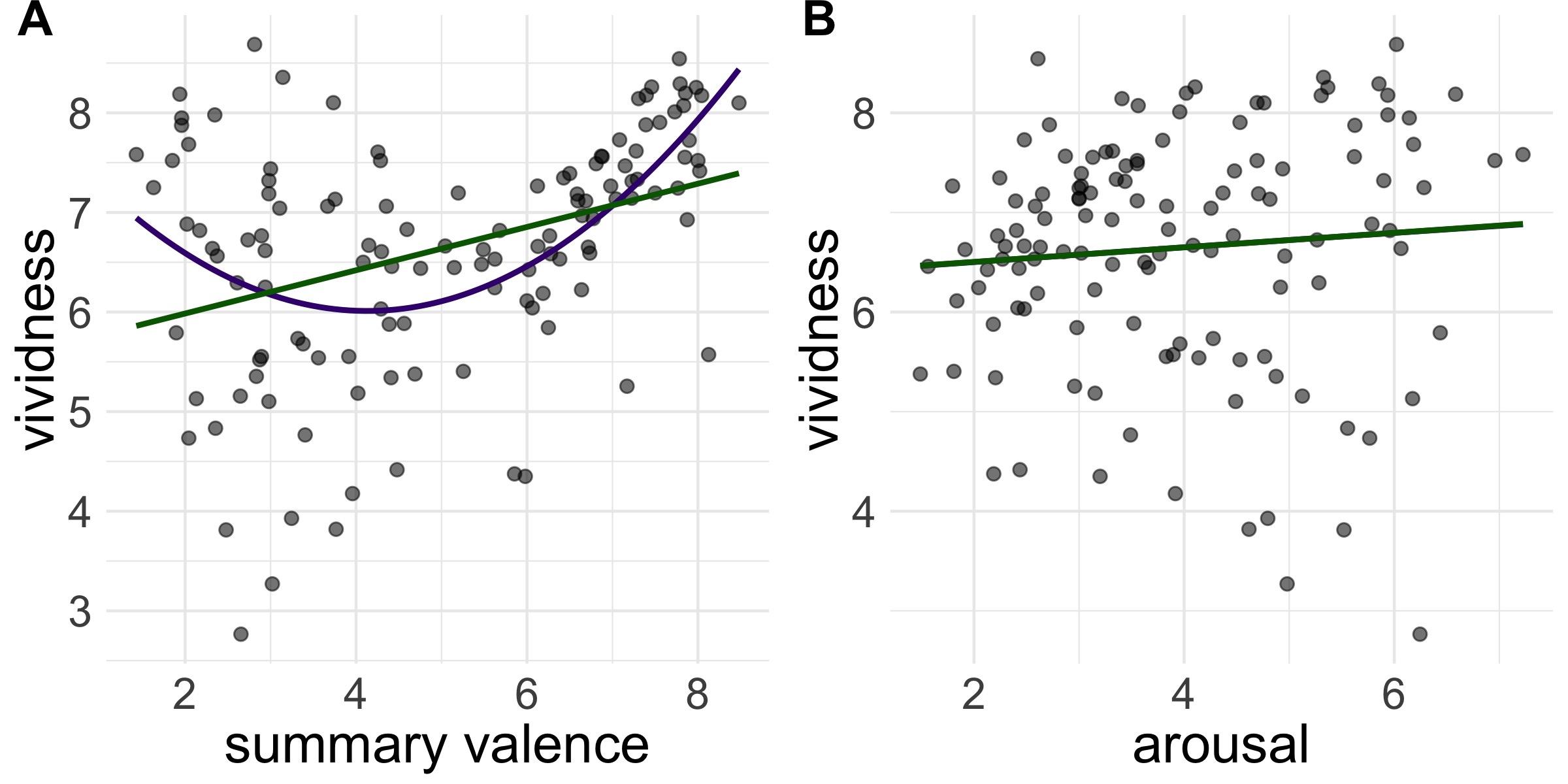
*Subjective memory vividness*. We first analyzed the relationship between visual and auditory memory vividness ratings and then tested the influence of summary valence and arousal on subsequent memory vividness. Visual vividness and auditory vividness were highly correlated (r = 0.96, *p* < .001) and so subsequent analyses use a single composite vividness measure calculated as the average of visual and auditory vividness on a trial-by-trial basis (see Table 1). To examine the relationship between memory vividness, summary valence and arousal, we tested a linear mixed effects model with vividness as the dependent variable and fixed effects terms representing linear and quadratic effects of summary valence and arousal. We allowed subject and video intercepts to vary as random effects. Summary valence influenced vividness linearly ((1) = 26.63, *p* < .001), 𝛽 -0.64 *±* 0.05, indicating that positive videos were remembered more vividly than more neutral or negative videos. The quadratic summary valence term was also significant ((1) = 61.86, *p* < .001), 𝛽 = 0.04 *±* 0.005, meaning that highly negative or highly positive videos were remembered more vividly than more neutral videos. Neither the linear nor the quadratic arousal terms uniquely contributed to the model fit.

As with summary valence and arousal, we examined the stability of vividness by iteratively correlating each participant’s vector of composite vividness ratings across videos with the composite vividness ratings of every other participant who had watched the same videos. The mean of these correlations revealed stability of video memorability across subjects, r = 0.35.

*Table 1***.** Summary statistics for emotion and memory ratings across videos.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | **valence** | **arousal** | **visual vividness** | **auditory vividness** |
| **mean** | 5.03 | 3.92 | 6.94 | 6.34 |
| **sd** | 2.03 | 1.36 | 1.20 | 1.21 |
| **median** | 5.04 | 3.77 | 7.15 | 6.55 |
| **min** | 1.43 | 1.49 | 3.00 | 2.53 |
| **max** | 8.48 | 7.22 | 8.77 | 8.60 |

*Temporal memory precision*. Our final set of analyses focused on memory for video duration. To examine overall accuracy for duration estimation, the mean correlation between actual and estimated duration was calculated within each participant. Pearson’s correlation coefficients were averaged across participants after transformation to Fisher z values (mean z = 0.29, SE = 0.02). A one-sample t-test revealed that the true mean was significantly greater than zero, t(1,97) = 16.62, *p* < .001, meaning that participants were able to recall video durations with above-chance accuracy. We next asked whether temporal memory was influenced by vividness. To quantify memory precision for temporal duration, we calculated estimate error as (|*actual duration - duration estimate|) / actual duration* on a trial-by-trial basis. We then tested a linear mixed effects model with duration estimate error as the dependent variable and a fixed effects term representing linear effects of composite vividness. We allowed subject and video intercepts to vary as random effects. Vividness predicted duration error linearly ((1) = 69.41, *p* < .001), 𝛽 -0.01 *±* 0.001, indicating that the duration estimates of videos that were remembered vividly were most accurate (lower error). We next asked if valence or arousal modulated memory for video duration. Here, we tested a linear mixed effects model with duration estimate error as the dependent variable and six fixed effects terms representing linear and quadratic effects of summary valence, arousal, and composite vividness. As before, we allowed subject and video intercepts to vary as random effects. Neither the linear or quadratic terms of valence, arousal, or vividness significantly contributed to this model fit, indicating that vividness no longer significantly predicted duration error when valence and arousal were accounted for.



*Figure 4.* Each video plotted by mean ratings of (A) composite vividness (1= least vivid, 9 = most vivid) by valence (1 = most negative, 9 = most positive), and (B) composite vividness by arousal (1 = least intense, 9 = most intense). Green lines represent the linear relation between measures; purple lines represent the quadratic relationship between measures.

# DISCUSSION

Naturalistic video stimuli have been increasingly used to investigate the cognitive and neural processes involved in representing dynamic events. Although emotion is a natural part of real experiences and can strongly influence event processing and memory, previous studies have not systematically manipulated the emotional content of such videos. In this study, we developed a new stimulus set of real news broadcasts and collected ratings of emotionality and subjective memory for each video. The results of our norming study indicated that, as intended, the stimulus set varied in its emotional content and memorability, and that the emotional valence predicted the subjective vividness with which the videos were remembered.

In this study, we collected summary ratings of valence and arousal as well as continuous ratings of valence. The summary ratings showed a pattern similar to that reported for other kinds of emotional stimulus sets (e.g., Marchewka et al., 2014): valence and arousal were significantly related to one another. There was a linear relationship reflecting that videos that were more emotionally negative tended to be most emotionally intense, consistent with past work documenting a negativity bias for participant ratings of the valence and intensity of emotional stimuli (Kuppens et al., 2013). There was also a quadratic (V-shaped) relationship, reflecting that arousal was greatest for the most negative and most positive videos, replicating many prior studies of these emotional components (e.g., Lang & Bradley, 2007).

Because of the dynamic nature of the stimuli, we were especially interested in the properties of the continuous valence ratings, which reflect the changing emotional experiences of the participants while they viewed the videos. The continuous ratings were strongly related to the summary valence ratings, both in terms of their average rating as well as their absolute peak (negative or positive) rating. We also found that arousal was related to the standard deviation of continuous valence ratings. That is, the videos in which the continuous valence ratings changed the most were the most arousing videos. Interestingly, this was true even when the first 5 seconds of the video were excluded, and thus this relationship is unlikely to merely reflect the initial shift in ratings away from a neutral baseline. Rather, this pattern may reflect the gradual accumulation of information that influences people’s impression of emotional videos, leading to changes in valence ratings that are protracted across the entire video period. Alternatively, large changes in valence may have been unexpected and thus amplified emotional arousal.

Though our analyses focused on relationships between measures, this study also provided some interesting qualitative observations. For example, the video containing the largest deviation in dynamic valence is also the video remembered most vividly. This particular clip showed a young girl standing on a toilet, which initially seems humorous. However, it is later revealed that she is demonstrating what she was taught to do during an active shooter drill at her preschool. The video with the most negative mean summary valence was also the video with the most negative mean peak: a report that white police officers who had beaten a black motorcyclist were acquitted by an all-white jury. Likewise, the video with the most positive mean summary valence also had the most positive mean peak: footage of an excited dog greeting a soldier who had just returned home. This exemplifies the strong relationship between peak valence and summary valence. Overall, though, participants rated the videos as varying widely along spectrums of valence, which makes them ideal for studies examining differential effects of positive and negative emotion. Because valence and arousal were rated as separate measures, these videos can be used to examine differential effects of these variables.

A common finding in the memory literature is that emotional memories are more subjectively vivid than their neutral counterparts (for review see LaBar & Cabeza, 2006). We sought to replicate this finding with the current stimulus set by cuing participants with the first 3 seconds of the video and having them rate the subjective vividness of their memories for the cued video. We found that summary valence ratings were related to memory vividness both linearly and quadratically. This pattern indicates that highly emotional videos were remembered more vividly than neutral videos, and memory vividness was particularly high for positive videos. Laboratory experiments using words or static images typically find memory enhancements for negative items (for review see Kensinger, 2007), but enhancements in memory for positive materials tend to be more variable (for review see Bennion et al., 2013), although they are often seen for autobiographical events (D’Argembeau et al., 2003). The current stimulus set may be particularly useful, then, for investigating the effects of valence on memory, as well as for relating studies of laboratory-based and autobiographical memories. Since memory was not objectively tested for details other than video duration, future work on emotional enhancement should examine what aspects of the videos receive the greatest mnemonic benefits.

Our video stimuli build upon existing video stimulus databases in several ways. Firstly, this set includes more stimuli than comparable databases. This is particularly important for experiments that plan to examine a subset of the data - for example, a memory experiment that compares remembered items between valence categories. Additionally, each video can be treated as a discrete event or videos can be played sequentially like a news broadcast to create a longer event that is emotionally variable yet thematically cohesive. Finally, this video stimuli set is, to our knowledge, the first of its kind to provide dynamic valence ratings of real-world emotional events. As such, it can facilitate examination of the neural and cognitive processes underlying dynamic, complex experiences. For example, it is unclear how changes in emotionality are related to later memory accuracy and vividness. In other words, do ‘spikes’ in emotion enhance memory for only those emotion-eliciting parts of an event, or do moments of emotional transition (e.g., an emotionally negative moment becoming positive) lead to better memory? Is memory for less emotional parts of an event also enhanced, or are neutral moments overshadowed by more salient information and forgotten? In addition to questions about the direct effects of emotion, these stimuli can be used to understand how we can effectively modulate our emotions. For example, emotion regulation strategies have differential effects on both experiential negative emotion and memory for an emotional event (Richards & Gross, 2000), as well as long-term mental health (Gross & John, 2003). However, the neural mechanisms that underlie healthy versus unhealthy emotion regulation are largely unknown. Answering questions such as these is critical to fully understanding the dynamic effects of emotion and how we can modulate these effects.

# Acknowledgments

This work was supported in part by NIH R00MH103401 grant (M.R.) and the Brain & Behavior Research Foundation NARSAD Young Investigator Grant (M.R.). We thank Cayley Bliss, Maximilian Bluestone, Emily Iannazzi, and Elizabeth Judge for their assistance with stimuli and data collection. The authors report no conflicts of interest.

# References

Abadi, M. K., Subramanian, R., Kia, S. M., Avesani, P., Patras, I., & Sebe, N. (2015). DECAF: MEG-Based Multimodal Database for Decoding Affective Physiological Responses. *IEEE Transactions on Affective Computing*, *6*(3), 209–222.

Abraham, A., von Cramon, D. Y., & Schubotz, R. I. (2008). Meeting George Bush versus meeting Cinderella: the neural response when telling apart what is real from what is fictional in the context of our reality. *Journal of Cognitive Neuroscience*, *20*(6), 965–976.

Bainbridge, W. A. (2017). The memorability of people: Intrinsic memorability across transformations of a person’s face. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *43*(5), 706–716.

Baldassano, C., Chen, J., Zadbood, A., Pillow, J. W., Hasson, U., & Norman, K. A. (2017). Discovering Event Structure in Continuous Narrative Perception and Memory. *Neuron*, *95*(3), 709–721.e5.

Bates, D., Maechler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. Journal of Statistical Software, 67(1), 1-48.<doi:10.18637/jss.v067.i01>.

Bennion, K. A., Ford, J. H., Murray, B. D., & Kensinger, E. A. (2013). Oversimplification in the study of emotional memory. *Journal of the International Neuropsychological Society: JINS*, *19*(9), 953–961.

Bird, C. M., Keidel, J. L., Ing, L. P., Horner, A. J., & Burgess, N. (2015). Consolidation of Complex Events via Reinstatement in Posterior Cingulate Cortex. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *35*(43), 14426–14434.

Bisby, J. A., & Burgess, N. (2013). Negative affect impairs associative memory but not item memory. *Learning & Memory* , *21*(1), 21–27.

Bisby, J. A., Horner, A. J., Hørlyck, L. D., & Burgess, N. (2016). Opposing effects of negative emotion on amygdalar and hippocampal memory for items and associations. *Social Cognitive and Affective Neuroscience*, *11*(6), 981–990.

Bradley, M. M., & Lang, P. J. (1999). *Affective norms for English words (ANEW): Instruction manual and affective ratings*. Citeseer. Retrieved from http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.306.3881&rep=rep1&type=pdf

Chen, J., Leong, Y. C., Honey, C. J., Yong, C. H., Norman, K. A., & Hasson, U. (2017). Shared memories reveal shared structure in neural activity across individuals. Nature Neuroscience (20)1, 115-25. doi:10.1038/nn.4450, http://arks.princeton.edu/ark:/88435/dsp01nz8062179

Cheng, R.-K., Tipples, J., Narayanan, N. S., & Meck, W. H. (2016). Clock Speed as a Window into Dopaminergic Control of Emotion and Time Perception. *Timing & Time Perception*, *4*(1), 99–122.

Curio, C., Bulthoff, H. H., and Giese, M. A. (2010). Dynamic Faces: Insights from Experiments and Computation.Cambridge, MA: MIT Press.

D’Argembeau, A., Comblain, C., & Van der Linden, M. (2003). Phenomenal characteristics of autobiographical memories for positive, negative, and neutral events. *Applied Cognitive Psychology*, *17*(3), 281–294.

Ekman, P. (1984). Expression and the nature of emotion. In K. R. Scherer & P. Ekman, (Eds.), *Approaches to Emotion* (319–344). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Gross, J. J., & John, O. P. (2003). Individual differences in two emotion regulation processes: implications for affect, relationships, and well-being. *Journal of Personality and Social Psychology*, *85*(2), 348–362.

Kark, S. M., & Kensinger, E. A. (2015). Effect of emotional valence on retrieval-related recapitulation of encoding activity in the ventral visual stream. *Neuropsychologia*, *78*, 221–230.

Kensinger, E. A. (2007). Negative Emotion Enhances Memory Accuracy: Behavioral and Neuroimaging Evidence. *Current Directions in Psychological Science*, *16*(4), 213–218.

Kensinger, E. A. (2009). Remembering the Details: Effects of Emotion. *Emotion Review*. https://doi.org/10.1177/1754073908100432

Kensinger, E. A., Garoff-Eaton, R. J., & Schacter, D. L. (2007). How negative emotion enhances the visual specificity of a memory. *Journal of Cognitive Neuroscience*, *19*(11), 1872–1887.

Klin, A., Jones, W., Schultz, R., Volkmar, F., & Cohen, D. (2002). Visual fixation patterns during viewing of naturalistic social situations as predictors of social competence in individuals with autism. Arch. Gen. Psychiatry 59, 809–816.

Koelstra, S., Muhl, C., Soleymani, M., Lee, J., Yazdani, A., Ebrahimi, T., Pun, T., Nijholt, A., & Patras, I. (2012). DEAP: A Database for Emotion Analysis ;Using Physiological Signals. *IEEE Transactions on Affective Computing*, *3*(1), 18–31.

Kuhlmann, S., Piel, M., & Wolf, O. T. (2005). Impaired memory retrieval after psychosocial stress in healthy young men. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *25*(11), 2977–2982.

Kuppens, P., Tuerlinckx, F., Russell, J. A., & Barrett, L. F. (2013). The relation between valence and arousal in subjective experience. *Psychological Bulletin*, *139*(4), 917–940.

LaBar, K. S., & Cabeza, R. (2006). Cognitive neuroscience of emotional memory. *Nature Reviews. Neuroscience*, *7*(1), 54–64.

Lang, P. J., & Bradley, M. M. (2007). The International Affective Picture System (IAPS) in the study of emotion and attention. In J.A. Coan & J.J.B. Allen (Eds.), *Handbook of Emotion Elicitation and Assessment* (29-46). New York, New York: Oxford University Press.

Lang, P. J., Bradley, M. M., & Cuthbert, B.N. (2008). International affective picture system (IAPS): Affective ratings of pictures and instruction manual. Technical Report A-8. University of Florida, Gainesville, FL.

Lang, P. J., Bradley, M. M., Fitzsimmons, J. R., Cuthbert, B. N., Scott, J. D., Moulder, B., & Nangia, V. (1998). Emotional arousal and activation of the visual cortex: an fMRI analysis. *Psychophysiology*, *35*(2), 199–210.

Lee, T.-H., Baek, J., Lu, Z.-L., & Mather, M. (2014). How arousal modulates the visual contrast sensitivity function. *Emotion*, *14*(5), 978–984.

Mackinnon, A., Jorm, A. F., Christensen, H., Korten, A. E., Jacomb, P. A., & Rodgers, B. (1999). A short form of the Positive and Negative Affect Schedule: evaluation of factorial validity and invariance across demographic variables in a community sample. *Personality and Individual Differences*, *27*(3), 405–416.

Marchewka, A., Zurawski, Ł., Jednoróg, K., & Grabowska, A. (2014). The Nencki Affective Picture System (NAPS): introduction to a novel, standardized, wide-range, high-quality, realistic picture database. *Behavior Research Methods*, *46*(2), 596–610.

Noulhiane, M., Mella, N., Samson, S., Ragot, R., & Pouthas, V. (2007). How emotional auditory stimuli modulate time perception. *Emotion*, *7*(4), 697.

Oedekoven, C. S. H., Keidel, J. L., Berens, S. C., & Bird, C. M. (2017). Reinstatement of memory representations for lifelike events over the course of a week. *Scientific Reports*, *7*(1), 14305.

Phelps, E. A., Ling, S., & Carrasco, M. (2006). Emotion facilitates perception and potentiates the perceptual benefits of attention. *Psychological Science*, *17*(4), 292–299.

Phelps, E. A., & Sharot, T. (2008). How (and Why) Emotion Enhances the Subjective Sense of Recollection. *Current Directions in Psychological Science*, *17*(2), 147–152.

Putman, P., Hermans, E., & van Honk, J. (2006). Anxiety meets fear in perception of dynamic expressive gaze. *Emotion* , *6*(1), 94–102.

R Core Team (2018). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

Richards, J. M., & Gross, J. J. (2000). Emotion regulation and memory: the cognitive costs of keeping one’s cool. *Journal of Personality and Social Psychology*, *79*(3), 410–424.

Richell, R. A., & Anderson, M. (2004). Reproducibility of negative mood induction: a self-referent plus musical mood induction procedure and a controllable/uncontrollable stress paradigm. *Journal of Psychopharmacology*, *18*(1), 94–101.

Risko, E. F., Laidlaw, K., Freeth, M., Foulsham, T., & Kingstone, A. (2012). Social attention with real versus reel stimuli: toward an empirical approach to concerns about ecological validity. *Frontiers in Human Neuroscience*, *6*, 143.

RStudio Team (2016). RStudio: Integrated Development for R. RStudio, Inc., Boston, MA URL http://www.rstudio.com/.

Schaefer, A., Nils, F., Sanchez, X., & Philippot, P. (2010). Assessing the effectiveness of a large database of emotion-eliciting films: A new tool for emotion researchers. *Cognition and Emotion* 24(7): 1153-1172. Retrieved from https://sites.uclouvain.be/ipsp/FilmStim/

Soleymani, M., Lichtenauer, J., Pun, T., & Pantic, M. (2012). A Multimodal Database for Affect Recognition and Implicit Tagging. *IEEE Transactions on Affective Computing*, *3*(1), 42–55.

Stroud, L. R., Tanofsky-Kraff, M., Wilfley, D. E., & Salovey, P. (2000). The Yale Interpersonal Stressor (YIPS): affective, physiological, and behavioral responses to a novel interpersonal rejection paradigm. *Annals of Behavioral Medicine: A Publication of the Society of Behavioral Medicine*, *22*(3), 204–213.

Tipples, J. (2008). Negative emotionality influences the effects of emotion on time perception. *Emotion*, *8*(1), 127–131.

Xie, W., & Zhang, W. (2017). Negative emotion enhances mnemonic precision and subjective feelings of remembering in visual long-term memory. *Cognition*, *166*, 73–83.

Yonelinas, A. P., & Ritchey, M. (2015). The slow forgetting of emotional episodic memories: an emotional binding account. *Trends in Cognitive Sciences*, *19*(5), 259–267.

Zadra, J. R., & Clore, G. L. (2011). Emotion and perception: the role of affective information. *Wiley Interdisciplinary Reviews. Cognitive Science*, *2*(6), 676–685.