Neuron



# A human colliculus-pulvinar-amygdala pathway encodes negative emotion

## Highlights

- Multivariate fMRI models identify the colliculus-pulvinar-amygdala pathway in humans
- The pathway responds to negative visual and auditory but not painful stimuli
- The pathway is associated with broad patterns of cortico-amygdalar connectivity


## Authors

Philip A. Kragel, Marta Čeko, Jordan Theriault, ..., Martin A. Lindquist, Lisa Feldman Barrett, Tor D. Wager

## Correspondence

pkragel@emory.edu (P.A.K.), tor.d.wager@dartmouth.edu (T.D.W.)



## In brief

Kragel et al. implement a new functional connectivity technique that estimates inter-region connections using local multivariate patterns. They identify a superior colliculus-pulvinar-amygdala pathway at 7 T and 3 T that responds to aversive images and sounds but not painful stimuli. This method provides a general framework for identifying and characterizing functional pathways.


Kragel et al., 2021, Neuron *109*, 1–9
August 4, 2021 
https://doi.org/10.1016/j.neuron.2021.06.001

CellPress

Neuron

CellPress

Report

# A human colliculus-pulvinar-amygdala pathway encodes negative emotion

Philip A. Kragel,[1,2,3,9,10,*] Marta Čeko,[1] Jordan Theriault,[4] Danlei Chen,[4] Ajay B. Satpute,[4,5] Lawrence W. Wald,[5,6] Martin A. Lindquist,[7] Lisa Feldman Barrett,[4,5,6] and Tor D. Wager[1,8,*]

[1]Institute of Cognitive Science, University of Colorado Boulder, Boulder, CO 80309, USA
[2]Department of Psychology, Emory University, Atlanta, GA 30322, USA
[3]Department of Psychiatry and Behavioral Science, Emory University, Atlanta, GA 30322, USA
[4]Department of Psychology, Northeastern University, Boston, MA 02115, USA
[5]Athinoula A. Martinos Center for Biomedical Imaging, Boston, MA 02129, USA
[6]Harvard Medical School, Boston, MA 02129, USA
[7]Department of Biostatistics, Johns Hopkins University, Baltimore, MD 21205, USA
[8]Department of Psychological and Brain Sciences, Dartmouth College, Hanover, NH 03755, USA
[9]Present address: Department of Psychology, Emory University, Atlanta, GA 30322, USA
[10]Lead contact
*Correspondence: pkragel@emory.edu (P.A.K.), tor.d.wager@dartmouth.edu (T.D.W.)
https://doi.org/10.1016/j.neuron.2021.06.001

## SUMMARY

Animals must rapidly respond to threats to survive. In rodents, threat-related signals are processed through a subcortical pathway from the superior colliculus to the amygdala, a putative "low road" to affective behavior. This pathway has not been well characterized in humans. We developed a novel pathway identification framework that uses pattern recognition to identify connected neural populations and optimize measurement of inter-region connectivity. We first verified that the model identifies known thalamocortical pathways with high sensitivity and specificity in 7 T (n = 56) and 3 T (n = 48) fMRI experiments. Then we identified a human functional superior colliculus-pulvinar-amygdala pathway. Activity in this pathway encodes the intensity of normative emotional responses to negative images and sounds but not pleasant images or painful stimuli. These results provide a functional description of a human "low road" pathway selective for negative exteroceptive events and demonstrate a promising method for characterizing human functional brain pathways.

## INTRODUCTION

Rapidly detecting threats in the environment is crucial for survival. In rodents, behavioral responses to visual danger signals are mediated by neural pathways from the superior colliculus to the pulvinar (Zhou et al., 2017) and from the pulvinar to the amygdala (Wei et al., 2015). This pathway, sometimes referred to as the "low road" to amygdala-mediated threat (de Gelder et al., 2011; Ledoux, 1998; Pessoa and Adolphs, 2010), enables reflexive behavioral responses (e.g., orienting) via collicular projections (Schneider, 1969) and interfaces with mechanisms in the amygdala that shape current and future defensive behaviors (Tamietto and de Gelder, 2010). The targets of this pathway, neurons in the basolateral amygdala (Wei et al., 2015), comprise one of many interdigitated neural populations involved in pain-related affect (Corder et al., 2019), anxiety (Tye et al., 2011), and reward and punishment (Burgos-Robles et al., 2017), among other behaviors (Janak and Tye, 2015). In humans, this pathway is thought to mediate unconscious processing of affective visual stimuli (Morris et al., 1999; Tamietto and de Gelder, 2010; Vuilleumier et al., 2003; Whalen et al., 2004). However, the pathways identified in animal research reflect communication among specific neural ensembles (Janak and Tye, 2015; Kyriazi et al., 2018) that are not clearly resolved with human neuroimaging. Although structural connections consistent with the colliculus-pulvinar-amygdala pathway have been identified (Abivardi and Bach, 2017; McFadyen et al., 2019; Rafal et al., 2015; Tamietto et al., 2012), its precise location and functions in humans require further study. In particular, it has not been functionally distinguished from other amygdalar circuits, and there is so far no defined measure of functional connectivity in this pathway. More precise identification of the pathway would allow more accurate characterization of its relationships with human affect and behavior.

Adoption of pattern recognition techniques has moved fMRI closer to the level of neural representation in humans (Kragel et al., 2018; Norman et al., 2006; Poldrack and Farah, 2015), potentially even uncovering information coded by cortical columns (Haynes and Rees, 2005; Kamitani and Tong, 2005), but the mainstay of connectivity research in humans examines relationships between coarse anatomical regions, not neural populations. As a result, there are ongoing debates about the

Please cite this article in press as: Kragel et al., A human colliculus-pulvinar-amygdala pathway encodes negative emotion, Neuron (2021), https://doi.org/10.1016/j.neuron.2021.06.001



colliculus-pulvinar-amygdala pathway, including whether it exists in primates (Pessoa and Adolphs, 2010). If so, it is unclear whether it selectively responds to emotional events (as opposed to attentionally demanding or behaviorally relevant events more generally) and whether it relates to human subjective emotional experience (LeDoux, 2014). We addressed these questions by extending pattern recognition approaches in fMRI (Haxby et al., 2001; Kriegeskorte et al., 2006; Poldrack and Farah, 2015) to modeling multivariate brain pathways in humans. We sought to identify a human colliculus-pulvinar-amygdala pathway and characterize its relationships with human emotion and pain.

Recent advances in connectivity models using multivariate fMRI patterns have the potential to identify human pathways connecting specific neural populations with increased precision (Anzellotti and Coutanche, 2018; Anzellotti et al., 2017; Coutanche and Thompson-Schill, 2013; Woo et al., 2014). Conventional approaches focus on connectivity among anatomical regions or large functional parcels (Friston, 1994), which averages over many off-target signals. Multivariate pattern analysis can help by identifying signals that are more sensitive to particular subsets of neural populations (Haxby et al., 2001; Kamitani and Tong, 2005). For example, distinct multivariate patterns within the same brain region can exhibit unique profiles of connectivity and differential associations across varieties of affect (Anzellotti et al., 2017; Anzellotti and Coutanche, 2018; Basti et al., 2019; Coutanche and Thompson-Schill, 2013; Woo et al., 2014).

The approach we developed, multivariate pathway identification (MPathI), was designed specifically to model connections between distinct populations located in different brain regions, as demonstrated frequently in contemporary animal studies (Deisseroth, 2011). MPathI accomplishes this goal using an extension of partial least-squares to optimize the covariance between two multivariate patterns (latent sources reflective of neural populations) and estimate the strength of functional correlations between them. Training the model to identify a population-level (group) pattern and testing on independent participants (Kragel et al., 2018; Woo et al., 2017) stabilizes model weights, prioritizes generalizability, and provides unbiased estimates of (1) connectivity measured by the pathway model, (2) prediction of subjective emotion, and (3) differences from standard connectivity measures. Unlike similar measures (e.g., canonical correlation; Hardoon et al., 2007)), optimizing covariance prioritizes identification of larger, more robust signals in each region (STAR Methods). In addition to testing whether particular pathways can be identified in humans with the assistance of pattern recognition algorithms, MPathI and related techniques provide optimized measures of pathways that can be tested for relationships with behavior, psychopathology, and treatments.

## RESULTS

We first evaluated whether MPathI can be used to identify thalamocortical pathways with high sensitivity and specificity. This is important because fMRI-based measures of functional correlations may be driven by signals other than direct connections (e.g., by global signals or indirect connections). To this end, we acquired high-resolution (1.1 mm isotropic) fMRI at 7 T with whole-brain coverage (n = 56; STAR Methods). We estimated the strength of connectivity in monosynaptic pathways from the lateral geniculate nucleus (LGN) to the primary visual cortex (V1) and from the medial geniculate nucleus (MGN) to the primary auditory cortex (A1). Pathway models produced multivoxel patterns in each pair of regions that maximally covary (i.e., $LGN_{V1}$-$V1_{LGN}$ and $MGN_{A1}$-$A1_{MGN}$). We compared the strength of these connections against crossed (off-target) connections (i.e., $LGN_{V1}$-$V1_{MGN}$ and $MGN_{A1}$-$A1_{LGN}$), which served as negative controls. Connectivity in target pathways (i.e., $LGN_{V1}$-$V1_{MGN}$ and $MGN_{A1}$-$A1_{LGN}$) was strong and positive (Pearson's $r$ = 0.3891 ± 0.0167 SEM, $p < 0.0001$), whereas "crossed" connections were nonsignificant ($r$ = 0.0359 ± 0.0205 SEM). The direct comparison was highly significant ($\Delta r$ = 0.353 ± 0.024 SEM, $t_{55}$ = 13.83, $p < 0.0001$; Figure 1B). Connectivity strength discriminated between target and off-target pathways with 85% ± 2.4% accuracy (sensitivity = 95%, 95% confidence interval [CI] = [90% 98%]; specificity = 76% [68% 84%]; area under the receiver operating characteristic curve [AUROC] = 0.92).

These findings were replicated in a separate fMRI study at conventional 3 T field strength (experiment 2, n = 48), with some loss of precision but good sensitivity and specificity. Experiment 2 differed in other important ways as well; it included stimulation with affective visual, auditory, mechanical, and thermal stimuli, and the MPathI model was estimated at the population (group) level, allowing correlation between pathway response and affective experience (see below and STAR Methods). MPathI demonstrated positive cross-validated connectivity estimates for target pathways ($r$ = 0.547 ± 0.031 SEM, $p < 0.0001$) that were stronger than estimates for off-target connections ($r$ = 0.492 ± 0.027 SEM; $\Delta r$ = 0.055 ± 0.016 SEM, $t_{47}$ = 3.86, $p < 0.0001$; accuracy = 76.0%; AUROC = 0.783). These observations suggest that MPathI is sensitive to connections in known thalamocortical pathways and shows regional specificity across nearby regions of the thalamus. Such specificity would not be expected if pathway connectivity estimates were driven by global signals shared across multiple brain regions.

Given evidence that MPathI can accurately identify known pathways, we next examined whether it could be used to detect a superior colliculus-pulvinar-amygdala pathway in humans at 7 T. This pathway is quiescent at rest in non-human animals (Ramcharan et al., 2005; Zhou et al., 2018), making it difficult to isolate in the absence of behavior (e.g., during resting-state fMRI). To induce relevant activity, we scanned participants during tasks involving visual manipulation of negative emotion and painful mechanical stimulation (STAR Methods). Despite evidence of anatomical connectivity in humans (Abivardi and Bach, 2017; Rafal et al., 2015) the colliculus and amygdala connect to potentially different pulvinar populations (Pessoa and Adolphs, 2010), making it unclear whether the pathway is conserved across species, and if it is, how best to optimize a measure that captures its connectivity. To accommodate the possibility of direct and multisynaptic pathways through the pulvinar (Pessoa and Adolphs, 2010), we fit separate models for the superior colliculus-pulvinar and pulvinar-amygdala segments of the pathway (Figure 2B). In experiment 1, MPathI revealed patterns of activity in the superior colliculus optimized to covary with activity in the pulvinar (which

Please cite this article in press as: Kragel et al., A human colliculus-pulvinar-amygdala pathway encodes negative emotion, Neuron (2021), https://doi.org/10.1016/j.neuron.2021.06.001



CellPress

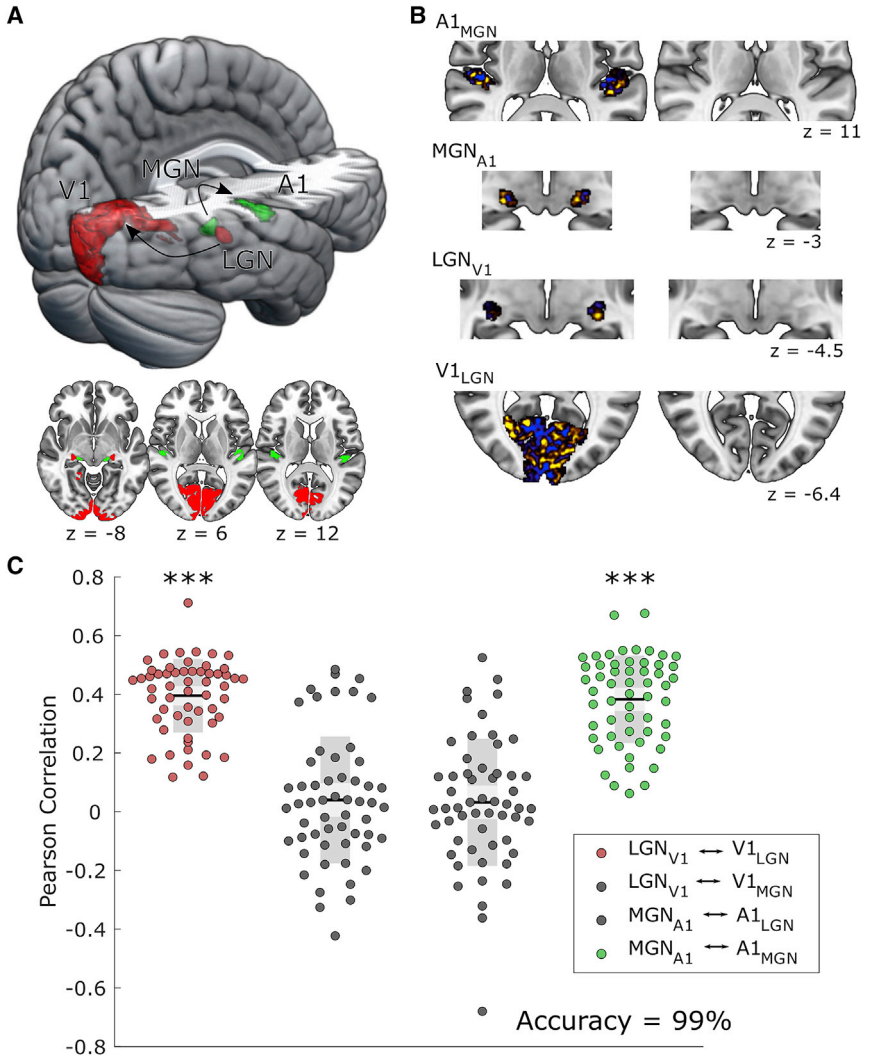


**Figure 1. Multivariate pathway identification (MPathI) discriminates direct from indirect thalamocortical sensory pathways**

(A) Visual (LGN-V1, red) and auditory (MGN-A1, green) regions of interest used to evaluate the sensitivity and specificity of MPathI.

(B) Axial views of patterns in the primary auditory cortex (A1), medial geniculate nucleus (MGN), lateral geniculate nucleus (LGN), and primary visual cortex (V1) that exhibit maximal functional connectivity with one another. Patterns were identified using partial least-squares to maximize connectivity between the MGN and A1 (i.e., patterns $MGN_{A1}$ and $A1_{MGN}$) in one model and LGN and V1 in another (i.e., patterns $LGN_{V1}$ and $V1_{LGN}$). Unthresholded patterns that exhibit maximal connectivity are shown for display purposes. Cool colors correspond to negative model weights, whereas positive colors indicate positive weights.

(C) Correlations for direct ($LGN_{V1}$-$V1_{LGN}$ and $MGN_{A1}$-$A1_{MGN}$) and indirect ($LGN_{V1}$-$V1_{MGN}$ and $MGN_{A1}$-$A1_{LGN}$) pathways fit on preprocessed BOLD time series from experiment 1 (n = 56). Each circle corresponds to one participant, solid black lines indicate mean, light gray regions indicate one standard deviation, and dark gray areas indicate 95% CIs. ***p < 0.0001.

we denote $SC_{Pulv}$) and patterns of activity in the pulvinar and amygdala optimized to covary with one another (i.e., $Pulv_{Amy}$ and $Amy_{Pulv}$; Figure 2B). The pulvinar-amygdala model ($r$ = 0.7071 ± 0.0179 SEM, $t_{55}$ = 39.588, p < 0.0001; Figure 2C) and superior colliculus-pulvinar model ($r$ = 0.7652 ± 0.0128 SEM, $t_{55}$ = 59.683, p < 0.0001) revealed strong connectivity. Connectivity was substantially stronger and more consistent across individuals than estimates from standard approaches based on region-average signals (pulvinar-amygdala: correlation difference $\Delta r$ = 0.402 ± 0.032 SEM, $t_{55}$ = 12.766, p < 0.0001; superior colliculus-pulvinar: $\Delta r$ = 0.328 ± 0.026 SEM, $t_{55}$ = 12.618, p < 0.0001). These observations demonstrate that our approach captures information conveyed by fine-scale patterns of fMRI activity that are not captured by standard connectivity measures.

In animal models, the superior colliculus-pulvinar-amygdala pathway is selectively activated by threat cues, particularly visual cues (Wei et al., 2015). In experiment 2 (n = 48), we evaluated whether human MPathI connectivity estimates are consistent with this pattern of selectivity. We scanned participants at 3 T during a multi-modal affect induction task designed to test the selectivity of affective responses. The task included five different types of stimuli assessed across two separate scanning sessions: negative images, negative sounds, mechanical pain, thermal pain, and positive images, with 4 intensity levels per stimulus type (20 event types). We estimated blood-oxygen-level-dependent (BOLD) responses to each stimulus presentation (6 trials × 4 intensity levels × 5 stimulus types × 48 participants, totaling 5,760 single-trial images). We fit pathway models using data aggregated across multiple subjects and evaluated their performance using cross-validation in independent subjects with 48-fold cross-validation. As in experiment 1, the pulvinar-amygdala model ($r$ = 0.683 ± 0.017 SEM, $t_{47}$ = 25.83, p < 0.0001) and superior colliculus-pulvinar model ($r$ = 0.859 ± 0.010 SEM, $t_{47}$ = 35.50, p < 0.0001) revealed strong connectivity (Figure 2C) at levels greater than those estimated using standard approaches based on region-average signals (pulvinar-amygdala: $\Delta r$ = 0.151 ± 0.021 SEM, $t_{47}$ = 8.22, p < 0.0001; superior colliculus-pulvinar: $\Delta r$ = 0.177 ± 0.012 SEM, $t_{47}$ = 16.97, p < 0.0001).

We next tested the colliculus-pulvinar-amygdala pathway's sensitivity and selectivity for negative emotional stimuli compared with salient painful and positive emotional stimuli. The colliculus-pulvinar-amygdala pathway in rodents conveys information about threats, whereas reward and pain signals reach the amygdala through different pathways; e.g., via other thalamic nuclei or parabrachial projections (Doron and Ledoux, 1999; Han et al., 2015; Palmiter, 2018; Tye et al., 2008).

Please cite this article in press as: Kragel et al., A human colliculus-pulvinar-amygdala pathway encodes negative emotion, Neuron (2021), https://doi.org/10.1016/j.neuron.2021.06.001



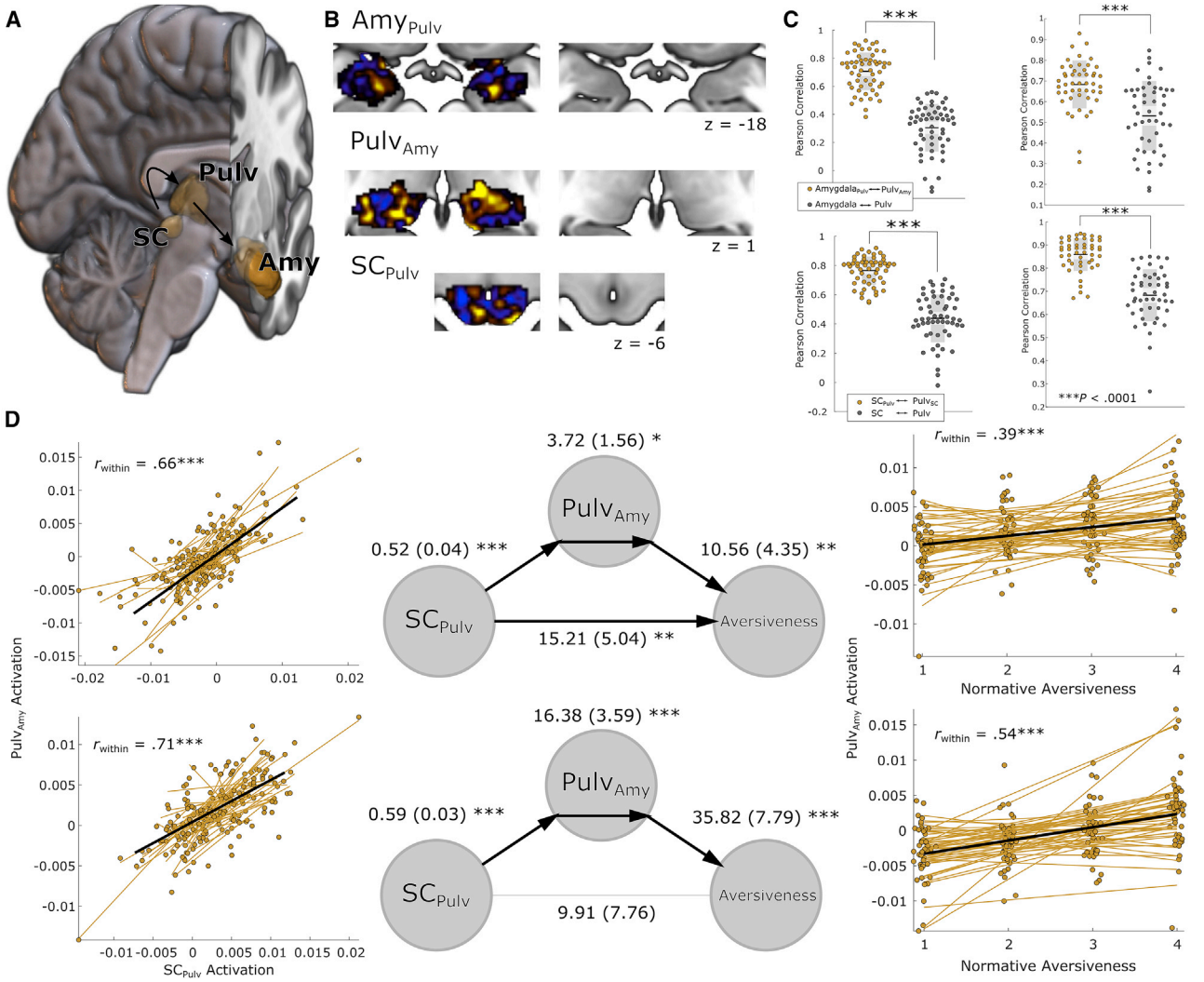


**Figure 2. A colliculus-pulvinar-amygdala pathway selective for negative emotion in humans**

(A) Hypothesized pathway and anatomical regions of interest rendered in MNI space.

(B) Axial views of patterns in the amygdala (Amy, top; peak $t_{55}$ = 3.36, p = 0.0014), pulvinar (Pulv, center; peak $t_{55}$ = 4.00, p = 0.0001), and superior colliculus (SC, bottom; peak $t_{55}$ = 3.32, p = .0016) that exhibit maximal functional connectivity with one another (target regions are denoted in subscript). Patterns that maximize connectivity between the Pulv and Amy were identified in one model and the SC and Pulv in another. Unthresholded patterns that exhibit maximal connectivity are shown for display purposes. Cool colors correspond to negative model weights, whereas positive colors indicate positive weights.

(C) Pathway models are more sensitive than conventional measures based on the mean signal of each region. Plots depict functional correlations estimated with multivariate pathway models (yellow circles) and mean signal in each region (gray circles). Each point corresponds to a single subject. The left panels depict data from experiment 1 (n = 56) and the right panels from experiment 2 (n = 48). Solid lines depict mean connectivity estimates, and boxes indicate standard deviation and 95% CIs.

(D) Activation of the Pulv$_{Amy}$ pattern mediates the relationship between SC$_{Pulv}$ activity and negative emotion for images (top row) and sounds (bottom row). Scatterplots show the relationship between SC$_{Pulv}$ activity and Pulv$_{Amy}$ activity (path *a*, left) and between Pulv$_{Amy}$ activity and negative emotion (path *b*, right). Lines show the least-squares fit between variables for each subject (n = 48). *p < 0.05, **p < 0.01, ***p < 0.001.

Therefore, we expected the pulvinar-amygdala pathway to selectively track the aversiveness of negative images (and possibly sounds) but not painful heat, painful pressure, or positive images. To test this prediction, we regressed Pulv$_{Amy}$ responses on normative rankings of aversiveness and pleasantness (see STAR Methods for details). Pulv$_{Amy}$ responses increased linearly with the subjective aversiveness of pictures ($\beta$ = 0.0011 ± 0.00026 SEM, $t_{47}$ = 4.363, p < 0.0001) and sounds ($\beta$ = 0.0019 ± 0.00026 SEM, $t_{47}$ = 4.363, p < 0.0001), showing that the pathway tracks negative emotion in humans across multiple sensory modalities. However, Pulv$_{Amy}$ responses showed little relationship to mechanical pain ($\beta$ = 0.0005 ± 0.00024 SEM, $t_{47}$ = 1.91, p = 0.0622), thermal pain ($\beta$ = 0.00025 ± 0.00029 SEM, $t_{47}$ = 0.852, p = 0.399), or pleasantness of positive pictures ($\beta$ = −0.00026 ± 0.00028 SEM, $t_{47}$ = −0.904, p = 0.371). Pulv$_{Amy}$ activation was more sensitive to the aversiveness of images and

Please cite this article in press as: Kragel et al., A human colliculus-pulvinar-amygdala pathway encodes negative emotion, Neuron (2021), https://doi.org/10.1016/j.neuron.2021.06.001





sounds than other stimulus types ($\Delta\beta$ = 0.0013 ± 0.00024 SEM, $t_{47}$ = 5.600, p < 0.0001). Similar results were found for colliculus and amygdala patterns (Table S1); however, interestingly, effect sizes were smallest in the amygdala, perhaps because of its functional heterogeneity (Janak and Tye, 2015; Kyriazi et al., 2018) or local fMRI signal inhomogeneity. These results establish that a pathway model optimized for colliculus-pulvinar-amygdala connectivity (without reference to behavior) predicts normative human judgements of aversiveness and that this pathway is sensitive and specific to exteroceptive processing of threatening stimuli as in non-human research (Wei et al., 2015).

Human affective responses may be mediated by direct and indirect pathways from the superior colliculus to the amygdala. The existence of a direct colliculus-pulvinar-amygdala pathway is contentious because threat-related collicular inputs target the inferior pulvinar, but pulvinar-amygdala projections originate in the medial pulvinar (Pessoa and Adolphs, 2010). This region exhibits strong functional correlations with higher-level visual areas (Arcaro et al., 2018), suggesting cortical involvement in threat processing rather than a direct subcortical pathway. However, other findings show that projections from inferior layers of the superior colliculus to the medial pulvinar and on to the amygdala mediate "fear-like" freezing in rodents (Wei et al., 2015). To evaluate evidence of a direct (versus indirect) pathway, we performed a multi-level mediation analysis (Atlas et al., 2010), testing whether fMRI activity in $Pulv_{Amy}$ (the pulvinar pattern optimized for amygdala connectivity) formally mediates the relationship between $SC_{Pulv}$ activity and normative human judgments of aversiveness. Significant mediation implies that $SC_{Pulv}$-$Pulv_{Amy}$ connectivity covaries with the $Pulv_{Amy}$ to emotion intensity connection and is most consistent with a direct pathway. We observed significant mediation effects (Figure 2D) for aversive pictures (path $a \times b$ = 3.720 ± 1.566 SEM, z = 2.356, p = 0.018) and sounds (path $a \times b$ = 16.379 ± 3.579 SEM, z = 3.580, p < 0.0001). A similar mediation analysis using online measures of self-reporting revealed similar effects for negative sounds (path $a \times b$ = 2.00 ± 0.450 SEM, z = 3.722, p < 0.0001), which varied primarily in terms of basic acoustic features but not images (path $a \times b$ = 0.2344 ± 0.2529 SEM, z = 0.8543, p = 0.3930), suggesting that this pathway may not be involved in evaluative processes that vary over time or between individuals. These results identify a human subcortical pathway involved in aversiveness and are consistent with direct projections from the superior colliculus to the pulvinar and on to the amygdala, but they do not preclude the existence of other pathways.

Standard connectivity approaches did not reveal significant mediation of aversiveness. Mediation using the mean signal in each region produced much smaller, non-significant effects for pictures (path $a \times b$ = 0.0029 ± 0.0171 SEM, z = 0.1728, p = 0.875, 95% CI of difference = [−2.738 to 13.964]) and sounds (path $a \times b$ = 0.0186 ± 0.0283 SEM, z = 0.711, p = 0.518, 95% CI of difference = [−13.890 to 31.820]). Thus, standard connectivity approaches are not sufficient to identify subcortical pathways mediating emotional experience.

The pathway we identified likely does not operate in isolation. Recent neuroscientific accounts (Pessoa and Adolphs, 2010) and studies of the amygdala's role in coordinating widely distributed activity (Gründemann et al., 2019; Stringer et al., 2019) suggest that the colliculus-pulvinar-amygdala pathway may be part of a broader cortical-subcortical network for generating and coordinating threat responses, and multiple cortico-amygdala pathways could play a similar role. To compare cortical and thalamic pathways with the amygdala, we used MPathI to identify optimized pathways connecting the amygdala with each of a series of local regions spanning the thalamus (Krauth et al., 2010) and cortex (Glasser et al., 2016). We performed principal-component analysis (PCA) on the amygdala-related cortical and thalamic pattern responses ("pathways" to the amygdala below) to identify commonalities in amygdala connectivity across multiple pathways (Figure 3) and situate the pulvinar-amygdala pathway among them.

This analysis revealed that activity in the pulvinar-amygdala pathway was functionally similar to pathways connecting the amygdala with multiple portions of "visual," "somatomotor," "frontoparietal," "attention," and "limbic" networks (Yeo et al., 2011). One group (component 1; Figure 3A) included the "low road" pulvinar-amygdala pathway, pathways from the insula and mid-cingulate cortex to the amygdala, and others. It was related to multiple large-scale networks (Figure 3B) and was the only component associated with "somatomotor," "attention," and "frontoparietal" networks. This group selectively responded to aversive images and sounds but not painful or positive stimuli (Figure 3C; Table S2). A second group of pathways (component 2) included parts of the traditional visual cortical "high road" to the amygdala, including amygdala connections with the LGN and visual cortex (predominantly the ventral stream) and was associated with "visual" and "limbic" cortical networks (Figure 3B). These pathways selectively responded to the aversiveness of negative pictures but not other stimuli. A third group (component 3) included amygdala connections with the anterior temporal lobe and orbitofrontal cortex (OFC) and loaded predominantly on the "limbic" cortical network (Figure 3B). These pathways were not associated with emotional intensity in this study. They are known to have dense connectivity with the amygdala and have been implicated in mnemonic (Olson et al., 2007) and associative (Shenhav et al., 2013) processing of affective information, including contextual influences on emotion that were not manipulated here. These findings identify three different profiles of amygdala activity, each correlated with distinct large-scale systems. The colliculus-pulvinar-amygdala pathway was part of a broader system anatomically and functionally consistent with a "low road" to visually and aurally induced threat, whereas other systems were consistent with "high road" visual cortical pathways to the amygdala and conceptual "top down" influences on the amygdala.

## DISCUSSION

Our findings are consistent with "multiple roads" accounts (Pessoa and Adolphs, 2010) but also consistent with the importance of the collicular-pulvinar-amygdala pathway for unconscious affective responses to visual stimuli (Tamietto and de Gelder, 2010). The amygdala appears to be important for some types of negative affect but not others, and other pathways may be more relevant for affective responses in other modalities (taste,

Please cite this article in press as: Kragel et al., A human colliculus-pulvinar-amygdala pathway encodes negative emotion, Neuron (2021), https://doi.org/10.1016/j.neuron.2021.06.001



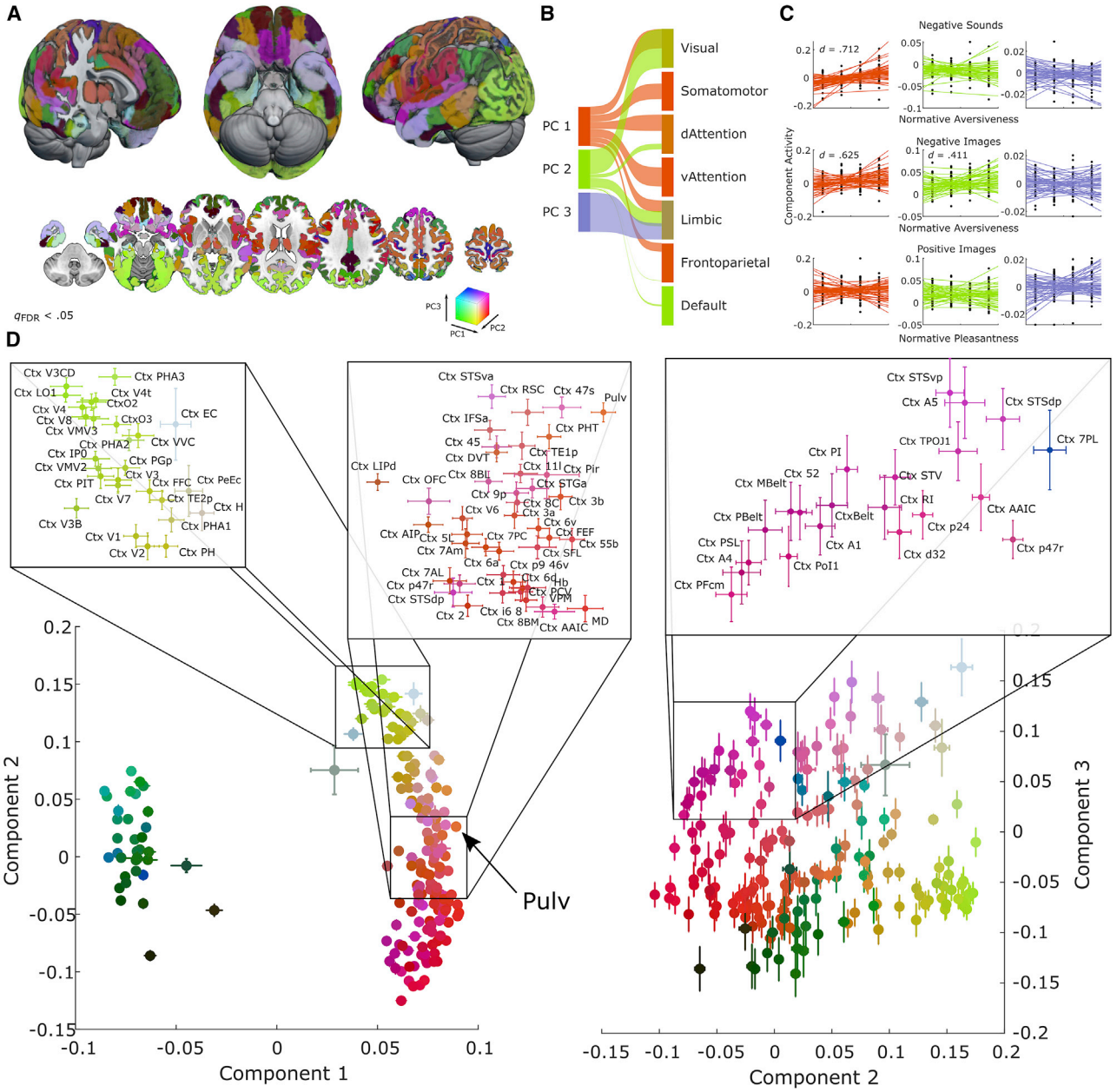


**Figure 3. Distributed pathways to the Amy are sensitive to negative stimuli**

(A) Decomposition of cortical and thalamic signals that covary with activity in the Amy, rendered on the brain. Pathway models were estimated between the Amy and cortical areas identified from a multimodal parcellation (Glasser et al., 2016) and an anatomical atlas of the thalamus and adjacent structures (17 regions). PCA reduces 197 estimates of activation that covary with the Amy to three orthogonal dimensions. Loadings of each region in this three-dimensional functional space are conveyed using an additive colorspace (red, component 1; green, component 2; blue, component 3).

(B) Spatial similarity of principal components and resting-state networks from Yeo et al. (2011). The width of each line indicates the degree of correlation between each component and binary maps of large-scale cortical networks. The first principal component is correlated with "somatomotor," "frontoparietal," and "attention" networks; the second component loads on "visual" and "limbic" networks; and the third component loads solely and predominantly on the "limbic" network.

(C) Relationship between activation of each component and the intensity of affective stimuli. Scatterplots show the relationship between activity of each component and the intensity of affective stimuli. Lines show the least-squares fit between variables for each subject (n = 48).

(D) Scatterplot depicting the mean and standard error of model coefficients for each region in the three-dimensional space, highlighting gradients of activity across regions (see Table S3 for full details). The Pulv region ($Pulv_{amy}$) of focus here is in the top right breakout, with some of the highest loadings on component 1 and near-zero loadings on components 2 and 3.

smell, or somatosensory). In addition, the relationship with conscious affective and emotional experience requires further study. Lesion studies indicate that the amygdala is not always necessary for emotional experience (Anderson and Phelps, 2002; Feinstein et al., 2013). We found that the colliculus-pulvinar pathway tracked normative affect responses across aversive sounds and images, demonstrating multi-modal properties and extending previous work showing pulvinar and amygdala coactivation in response to visual threat cues (Liddell et al., 2005; Morris et al., 1999; Vuilleumier et al., 2003). Pathway activity also tracked reported affective experience, but it did so more strongly for sounds than images. Notably, the sounds used here are perceptually simpler and more uniform than the emotional images, which may be evaluated variably across individuals and require more conceptual processing. Recent theories (Barrett, 2017; LeDoux and Brown, 2017) posit that transformation and re-representation of sensory signals is a defining feature of emotional experience and consciousness more generally (Dehaene et al., 2017). In light of these views and the many pathways that convey emotion-related information (Gothard, 2020), the sensory signals carried by the colliculus-pulvinar-amygdala pathway likely do not explain the full complexity of emotional experience in humans but contribute as one of multiple stages in a distributed information processing system.

This study provides more precise functional identification and characterization of the colliculus-pulvinar-amygdala pathway than what has been available previously in humans. The pathway exhibited a strong and selective relationship with the aversiveness of auditory and visual stimuli. This and other thalamocortical pathways were identified most clearly using ultra-high-field fMRI (7 T; see also Wang et al., 2020), demonstrating a benefit of high-resolution imaging that may provide increasing advantages as analytic tools develop to capitalize on it. However, the multivariate pattern approach inherent in MPathI helped to identify signals clearly related to specific functional pathways even at 3 T. MPathI measures substantially outperformed conventional region average-based connectivity and discriminated amygdala contributions from the colliculus-pulvinar pathway from at least two other components contributing to amygdala activity. These findings support the promise of MPathI and related techniques as important tools for "next-generation" human brain connectivity (Anzellotti and Coutanche, 2018; Anzellotti et al., 2017; Basti et al., 2020; Woo et al., 2017). The ability to identify such pathways in humans addresses a crucial gap between rapidly emerging animal research on neural pathways and assessment of functional correlations with human neuroimaging. It provides a framework for investigating the functional sensitivity and specificity of brain pathways across task paradigms and species and is a step toward understanding how activity in multiple pathways jointly relates to subjective affective experience.

## STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- RESOURCE AVAILABILITY
  - Lead contact
  - Materials availability
  - Data and code availability
- EXPERIMENTAL MODEL AND SUBJECT DETAILS
  - 7T fMRI Study (Expt 1)
  - 3T fMRI Study (Expt 2)
- METHOD DETAILS
  - 7T fMRI Study (Expt 1)
  - 3T fMRI Study (Expt. 2)
- QUANTIFICATION AND STATISTICAL ANALYSIS
  - MRI data acquisition and preprocessing
  - fMRI data analysis
  - Multivariate pathway identification
  - Sensitivity analysis
  - Control analyses
  - Mediation analysis
  - Cortical and thalamic pathway estimation

## SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.neuron.2021.06.001.

## ACKNOWLEDGMENTS

We thank Daniel Ott for assistance with data collection and Mickela Heilicher for literature review and thoughtful discussions about this project. This work was supported by National Institutes of Health grants R01 MH076136, DA046064, MH116026, and EB026549 and National Cancer Institute grant U01 CA193632.

## AUTHOR CONTRIBUTIONS

Conceptualization, methodology, formal analysis, writing – original draft, writing –review & editing, and visualization, P.A.K.; investigation, data curation, formal analysis, and writing – review & editing, M.Č.; investigation, data curation, and writing – reviewing & editing, J.T.; investigation, data curation, and writing – review & editing, D.C.; supervision, project administration, writing – review & editing, and funding acquisition, A.B.S.; supervision, writing – Review & editing, and funding acquisition, L.W.W.; conceptualization, methodology, writing – reviewing & editing, and funding acquisition, M.A.L.; conceptualization, supervision, project administration, writing – review & editing, and funding acquisition, L.F.B.; conceptualization, methodology, supervision, project administration, writing – original draft, review & editing, and funding acquisition, T.D.W.

## DECLARATION OF INTERESTS

The authors declare no competing interests.

Received: October 8, 2020
Revised: March 8, 2021
Accepted: June 1, 2021
Published: June 23, 2021

## SUPPORTING CITATIONS

The following reference appears in the Supplemental information: Keuken et al. (2014).

## REFERENCES

Abivardi, A., and Bach, D.R. (2017). Deconstructing white matter connectivity of human amygdala nuclei with thalamus and cortex subdivisions in vivo. Hum. Brain Mapp. *38*, 3927–3940.

Amunts, K., Kedo, O., Kindler, M., Pieperhoff, P., Mohlberg, H., Shah, N.J., Habel, U., Schneider, F., and Zilles, K. (2005). Cytoarchitectonic mapping of the human amygdala, hippocampal region and entorhinal cortex: intersubject variability and probability maps. Anat. Embryol. (Berl.) *210*, 343–352.

Anderson, A.K., and Phelps, E.A. (2002). Is the human amygdala critical for the subjective experience of emotion? Evidence of intact dispositional affect in patients with amygdala lesions. J. Cogn. Neurosci. *14*, 709–720.

Anzellotti, S., and Coutanche, M.N. (2018). Beyond Functional Connectivity: Investigating Networks of Multivariate Representations. Trends Cogn. Sci. *22*, 258–269.

Anzellotti, S., Caramazza, A., and Saxe, R. (2017). Multivariate pattern dependence. PLoS Comput. Biol. *13*, e1005799.

Arcaro, M.J., Pinsk, M.A., Chen, J., and Kastner, S. (2018). Organizing principles of pulvino-cortical functional coupling in humans. Nat. Commun. *9*, 5382.

Atlas, L.Y., Bolger, N., Lindquist, M.A., and Wager, T.D. (2010). Brain mediators of predictive cue effects on perceived pain. J. Neurosci. *30*, 12964–12977.

Avants, B.B., Epstein, C.L., Grossman, M., and Gee, J.C. (2008). Symmetric diffeomorphic image registration with cross-correlation: evaluating automated labeling of elderly and neurodegenerative brain. Med. Image Anal. *12*, 26–41.

Barrett, L.F. (2017). The theory of constructed emotion: an active inference account of interoception and categorization. Soc. Cogn. Affect. Neurosci. *12*, 1833.

Barron, D.S., Eickhoff, S.B., Clos, M., and Fox, P.T. (2015). Human pulvinar functional organization and connectivity. Hum. Brain Mapp. *36*, 2417–2431.

Basti, A., Mur, M., Kriegeskorte, N., Pizzella, V., Marzetti, L., and Hauk, O. (2019). Analysing linear multivariate pattern transformations in neuroimaging data. PLoS One *14*, e0223660.

Basti, A., Nili, H., Hauk, O., Marzetti, L., and Henson, R.N. (2020). Multi-dimensional connectivity: a conceptual and mathematical review. Neuroimage *221*, 117179.

Brainard, D.H. (1997). The Psychophysics Toolbox. Spat. Vis. *10*, 433–436.

Burgos-Robles, A., Kimchi, E.Y., Izadmehr, E.M., Porzenheim, M.J., Ramos-Guasp, W.A., Nieh, E.H., Felix-Ortiz, A.C., Namburi, P., Leppla, C.A., Presbrey, K.N., et al. (2017). Amygdala inputs to prefrontal cortex guide behavior amid conflicting cues of reward and punishment. Nat. Neurosci. *20*, 824–835.

Chang, J., and Yu, R. (2018). Alternations in functional connectivity of amygdalar subregions under acute social stress. Neurobiol. Stress *9*, 264–270.

Chang, L.J., Yarkoni, T., Khaw, M.W., and Sanfey, A.G. (2013). Decoding the role of the insula in human cognition: functional parcellation and large-scale reverse inference. Cereb. Cortex *23*, 739–749.

Corder, G., Ahanonu, B., Grewe, B.F., Wang, D., Schnitzer, M.J., and Scherrer, G. (2019). An amygdalar neural ensemble that encodes the unpleasantness of pain. Science *363*, 276–281.

Coutanche, M.N., and Thompson-Schill, S.L. (2013). Informational connectivity: identifying synchronized discriminability of multi-voxel patterns across the brain. Front. Hum. Neurosci. *7*, 15.

Cox, R.W. (1996). AFNI: software for analysis and visualization of functional magnetic resonance neuroimages. Comput. Biomed. Res. *29*, 162–173.

de Gelder, B., van Honk, J., and Tamietto, M. (2011). Emotion in the brain: of low roads, high roads and roads less travelled. Nat. Rev. Neurosci. *12*, 425–author reply 425.

Dehaene, S., Lau, H., and Kouider, S. (2017). What is consciousness, and could machines have it? Science *358*, 486–492.

Deisseroth, K. (2011). Optogenetics. Nat. Methods *8*, 26–29.

de Jong, S. (1993). SIMPLS: An alternative approach to partial least squares regression. Chemometrics Intellig. Lab. Syst. *18*, 251–263.

Doron, N.N., and Ledoux, J.E. (1999). Organization of projections to the lateral amygdala from auditory and visual areas of the thalamus in the rat. J. Comp. Neurol. *412*, 383–409.

Esteban, O., Markiewicz, C.J., Blair, R.W., Moodie, C.A., Isik, A.I., Erramuzpe, A., Kent, J.D., Goncalves, M., DuPre, E., Snyder, M., et al. (2019). fMRIPrep: a robust preprocessing pipeline for functional MRI. Nat. Methods *16*, 111–116.

Feinstein, J.S., Buzza, C., Hurlemann, R., Follmer, R.L., Dahdaleh, N.S., Coryell, W.H., Welsh, M.J., Tranel, D., and Wemmie, J.A. (2013). Fear and panic in humans with bilateral amygdala damage. Nat. Neurosci. *16*, 270–272.

Friston, K.J. (1994). Functional and effective connectivity in neuroimaging: A synthesis. Hum. Brain Mapp. *2*, 56–78.

Gattass, R., Soares, J.G.M., and Lima, B. (2017). The Pulvinar Thalamic Nucleus of Non-Human Primates: Architectonic and Functional Subdivisions (Springer).

Gazzaley, A., Rissman, J., and D'Esposito, M. (2004). Functional connectivity during working memory maintenance. Cogn. Affect. Behav. Neurosci. *4*, 580–599.

Geerligs, L., Cam-Can, and Henson, R.N. (2016). Functional connectivity and structural covariance between regions of interest can be measured more accurately using multivariate distance correlation. Neuroimage *135*, 16–31.

Glasser, M.F., Coalson, T.S., Robinson, E.C., Hacker, C.D., Harwell, J., Yacoub, E., Ugurbil, K., Andersson, J., Beckmann, C.F., Jenkinson, M., et al. (2016). A multi-modal parcellation of human cerebral cortex. Nature *536*, 171–178.

Gorgolewski, K., Burns, C.D., Madison, C., Clark, D., Halchenko, Y.O., Waskom, M.L., and Ghosh, S.S. (2011). Nipype: a flexible, lightweight and extensible neuroimaging data processing framework in python. Front. Neuroinform. *5*, 13.

Gothard, K.M. (2020). Multidimensional processing in the amygdala. Nat. Rev. Neurosci. *21*, 565–575.

Greve, D.N., and Fischl, B. (2009). Accurate and robust brain image alignment using boundary-based registration. Neuroimage *48*, 63–72.

Gründemann, J., Bitterman, Y., Lu, T., Krabbe, S., Grewe, B.F., Schnitzer, M.J., and Lüthi, A. (2019). Amygdala ensembles encode behavioral states. Science *364*, eaav8736.

Gu, X., Hof, P.R., Friston, K.J., and Fan, J. (2013). Anterior insular cortex and emotional awareness. J. Comp. Neurol. *521*, 3371–3388.

Guedj, C., and Vuilleumier, P. (2020). Functional connectivity fingerprints of the human pulvinar: Decoding its role in cognition. Neuroimage *221*, 117162.

Han, S., Soleiman, M.T., Soden, M.E., Zweifel, L.S., and Palmiter, R.D. (2015). Elucidating an Affective Pain Circuit that Creates a Threat Memory. Cell *162*, 363–374.

Hardoon, D.R., Mourão-Miranda, J., Brammer, M., and Shawe-Taylor, J. (2007). Unsupervised analysis of fMRI data using kernel canonical correlation. Neuroimage *37*, 1250–1259.

Haxby, J.V., Gobbini, M.I., Furey, M.L., Ishai, A., Schouten, J.L., and Pietrini, P. (2001). Distributed and overlapping representations of faces and objects in ventral temporal cortex. Science *293*, 2425–2430.

Haynes, J.-D., and Rees, G. (2005). Predicting the orientation of invisible stimuli from activity in human primary visual cortex. Nat. Neurosci. *8*, 686–691.

Hrybouski, S., Aghamohammadi-Sereshki, A., Madan, C.R., Shafer, A.T., Baron, C.A., Seres, P., Beaulieu, C., Olsen, F., and Malykhin, N.V. (2016). Amygdala subnuclei response and connectivity during emotional processing. Neuroimage *133*, 98–110.

Janak, P.H., and Tye, K.M. (2015). From circuits to behaviour in the amygdala. Nature *517*, 284–292.

Kamitani, Y., and Tong, F. (2005). Decoding the visual and subjective contents of the human brain. Nat. Neurosci. *8*, 679–685.

Keuken, M.C., Bazin, P.L., Crown, L., Hootsmans, J., Laufer, A., Müller-Axt, C., Sier, R., van der Putten, E.J., Schäfer, A., Turner, R., and Forstmann, B.U. (2014). Quantifying inter-individual anatomical variability in the subcortex using 7 T structural MRI. Neuroimage *94*, 40–46.

Kleckner, I.R., Zhang, J., Touroutoglou, A., Chanes, L., Xia, C., Simmons, W.K., Quigley, K.S., Dickerson, B.C., and Barrett, L.F. (2017). Evidence for a large-scale brain system supporting allostasis and interoception in humans. Nat. Hum. Behav. *1*, 1.

Kragel, P.A., Koban, L., Barrett, L.F., and Wager, T.D. (2018). Representation, Pattern Information, and Brain Signatures: From Neurons to Neuroimaging. Neuron *99*, 257–273.

Krauth, A., Blanc, R., Poveda, A., Jeanmonod, D., Morel, A., and Székely, G. (2010). A mean three-dimensional atlas of the human thalamus: generation from multiple histological data. Neuroimage *49*, 2053–2062.

Kriegeskorte, N., Goebel, R., and Bandettini, P. (2006). Information-based functional brain mapping. Proc. Natl. Acad. Sci. USA *103*, 3863–3868.

Kriegeskorte, N., Cusack, R., and Bandettini, P. (2010). How does an fMRI voxel sample the neuronal activity pattern: compact-kernel or complex spatiotemporal filter? Neuroimage *49*, 1965–1976.

Kumar, S., Forster, H.M., Bailey, P., and Griffiths, T.D. (2008). Mapping unpleasantness of sounds to their auditory representation. J. Acoust. Soc. Am. *124*, 3810–3817.

Kyriazi, P., Headley, D.B., and Pare, D. (2018). Multi-dimensional Coding by Basolateral Amygdala Neurons. Neuron *99*, 1315–1328.e5.

Lang, P.J., Bradley, M.M., and Cuthbert, B.N. (1997). International affective picture system (IAPS): Technical manual and affective ratings. https://www2.unifesp.br/dpsicobio/adap/instructions.pdf.

Le Floch, E., Guillemot, V., Frouin, V., Pinel, P., Lalanne, C., Trinchera, L., Tenenhaus, A., Moreno, A., Zilbovicius, M., Bourgeron, T., et al. (2012). Significant correlation between a set of genetic polymorphisms and a functional brain network revealed by feature selection and sparse Partial Least Squares. Neuroimage *63*, 11–24.

Ledoux, J. (1998). The Emotional Brain: The Mysterious Underpinnings of Emotional Life (Simon and Schuster).

LeDoux, J.E. (2014). Coming to terms with fear. Proc. Natl. Acad. Sci. USA *111*, 2871–2878.

LeDoux, J.E., and Brown, R. (2017). A higher-order theory of emotional consciousness. Proc. Natl. Acad. Sci. USA *114*, E2016–E2025.

Liddell, B.J., Brown, K.J., Kemp, A.H., Barton, M.J., Das, P., Peduto, A., Gordon, E., and Williams, L.M. (2005). A direct brainstem-amygdala-cortical 'alarm' system for subliminal signals of fear. Neuroimage *24*, 235–243.

May, P.J. (2006). The mammalian superior colliculus: laminar structure and connections. Prog. Brain Res. *151*, 321–378.

McFadyen, J., Mattingley, J.B., and Garrido, M.I. (2019). An afferent white matter pathway from the pulvinar to the amygdala facilitates fear recognition. eLife *8*, e40766.

Morris, J.S., Öhman, A., and Dolan, R.J. (1999). A subcortical pathway to the right amygdala mediating "unseen" fear. Proc. Natl. Acad. Sci. USA *96*, 1680–1685.

Norman, K.A., Polyn, S.M., Detre, G.J., and Haxby, J.V. (2006). Beyond mind-reading: multi-voxel pattern analysis of fMRI data. Trends Cogn. Sci. *10*, 424–430.

Olson, I.R., Plotzker, A., and Ezzyat, Y. (2007). The Enigmatic temporal pole: a review of findings on social and emotional processing. Brain *130*, 1718–1731.

Palmiter, R.D. (2018). The Parabrachial Nucleus: CGRP Neurons Function as a General Alarm. Trends Neurosci. *41*, 280–293.

Pessoa, L., and Adolphs, R. (2010). Emotion processing and the amygdala: from a 'low road' to 'many roads' of evaluating biological significance. Nat. Rev. Neurosci. *11*, 773–783.

Poldrack, R.A., and Farah, M.J. (2015). Progress and challenges in probing the human brain. Nature *526*, 371–379.

Rafal, R.D., Koller, K., Bultitude, J.H., Mullins, P., Ward, R., Mitchell, A.S., and Bell, A.H. (2015). Connectivity between the superior colliculus and the amygdala in humans and macaque monkeys: virtual dissection with probabilistic DTI tractography. J. Neurophysiol. *114*, 1947–1962.

Ramcharan, E.J., Gnadt, J.W., and Sherman, S.M. (2005). Higher-order thalamic relays burst more than first-order relays. Proc. Natl. Acad. Sci. USA *102*, 12236–12241.

Roy, M., Shohamy, D., Daw, N., Jepma, M., Wimmer, G.E., and Wager, T.D. (2014). Representation of aversive prediction errors in the human periaqueductal gray. Nat. Neurosci. *17*, 1607–1612.

Schneider, G.E. (1969). Two visual systems. Science *163*, 895–902.

Shenhav, A., Barrett, L.F., and Bar, M. (2013). Affective value and associative processing share a cortical substrate. Cogn. Affect. Behav. Neurosci. *13*, 46–59.

Smith, S.M., Nichols, T.E., Vidaurre, D., Winkler, A.M., Behrens, T.E.J., Glasser, M.F., Ugurbil, K., Barch, D.M., Van Essen, D.C., and Miller, K.L. (2015). A positive-negative mode of population covariation links brain connectivity, demographics and behavior. Nat. Neurosci. *18*, 1565–1567.

Stringer, C., Pachitariu, M., Steinmetz, N., Reddy, C.B., Carandini, M., and Harris, K.D. (2019). Spontaneous behaviors drive multidimensional, brainwide activity. Science *364*, 255.

Tamietto, M., and de Gelder, B. (2010). Neural bases of the non-conscious perception of emotional signals. Nat. Rev. Neurosci. *11*, 697–709.

Tamietto, M., Pullens, P., de Gelder, B., Weiskrantz, L., and Goebel, R. (2012). Subcortical connections to human amygdala and changes following destruction of the visual cortex. Curr. Biol. *22*, 1449–1455.

Tardif, E., Delacuisine, B., Probst, A., and Clarke, S. (2005). Intrinsic connectivity of human superior colliculus. Exp. Brain Res. *166*, 316–324.

Tye, K.M., Stuber, G.D., de Ridder, B., Bonci, A., and Janak, P.H. (2008). Rapid strengthening of thalamo-amygdala synapses mediates cue-reward learning. Nature *453*, 1253–1257.

Tye, K.M., Prakash, R., Kim, S.-Y., Fenno, L.E., Grosenick, L., Zarabi, H., Thompson, K.R., Gradinaru, V., Ramakrishnan, C., and Deisseroth, K. (2011). Amygdala circuitry mediating reversible and bidirectional control of anxiety. Nature *471*, 358–362.

Vuilleumier, P., Armony, J.L., Driver, J., and Dolan, R.J. (2003). Distinct spatial frequency sensitivities for processing faces and emotional expressions. Nat. Neurosci. *6*, 624–631.

Wang, Y.C., Bianciardi, M., Chanes, L., and Satpute, A.B. (2020). Ultra high field fMRI of human superior colliculi activity during affective visual processing. Sci. Rep. *10*, 1331.

Wei, P., Liu, N., Zhang, Z., Liu, X., Tang, Y., He, X., Wu, B., Zhou, Z., Liu, Y., Li, J., et al. (2015). Processing of visually evoked innate fear by a non-canonical thalamic pathway. Nat. Commun. *6*, 6756.

Whalen, P.J., Kagan, J., Cook, R.G., Davis, F.C., Kim, H., Polis, S., McLaren, D.G., Somerville, L.H., McLean, A.A., Maxwell, J.S., and Johnstone, T. (2004). Human amygdala responsivity to masked fearful eye whites. Science *306*, 2061.

Wold, S., Sjöström, M., and Eriksson, L. (2001). PLS-regression: a basic tool of chemometrics. Chemom. Intell. Lab. Syst. *58*, 109–130.

Woo, C.-W., Koban, L., Kross, E., Lindquist, M.A., Banich, M.T., Ruzic, L., Andrews-Hanna, J.R., and Wager, T.D. (2014). Separate neural representations for physical pain and social rejection. Nat. Commun. *5*, 5380.

Woo, C.-W., Chang, L.J., Lindquist, M.A., and Wager, T.D. (2017). Building better biomarkers: brain models in translational neuroimaging. Nat. Neurosci. *20*, 365–377.

Yeo, B.T.T., Krienen, F.M., Sepulcre, J., Sabuncu, M.R., Lashkari, D., Hollinshead, M., Roffman, J.L., Smoller, J.W., Zöllei, L., Polimeni, J.R., et al. (2011). The organization of the human cerebral cortex estimated by intrinsic functional connectivity. J. Neurophysiol. *106*, 1125–1165.

Zhang, Y., Brady, M., and Smith, S. (2001). Segmentation of brain MR images through a hidden Markov random field model and the expectation-maximization algorithm. IEEE Trans. Med. Imaging *20*, 45–57.

Zhou, N.A., Maire, P.S., Masterson, S.P., and Bickford, M.E. (2017). The mouse pulvinar nucleus: Organization of the tectorecipient zones. Vis. Neurosci. *34*, E011.

Zhou, N., Masterson, S.P., Damron, J.K., Guido, W., and Bickford, M.E. (2018). The Mouse Pulvinar Nucleus Links the Lateral Extrastriate Cortex, Striatum, and Amygdala. J. Neurosci. *38*, 347–362.

Please cite this article in press as: Kragel et al., A human colliculus-pulvinar-amygdala pathway encodes negative emotion, Neuron (2021), https://doi.org/10.1016/j.neuron.2021.06.001



# STAR★METHODS

## KEY RESOURCES TABLE

| REAGENT or RESOURCE | SOURCE | IDENTIFIER |
|---|---|---|
| Deposited data | | |
| Brain models and maps | This Paper | https://neurovault.org/collections/9982 |
| Processed fMRI data | This Paper | https://osf.io/werk2; https://doi.org/10.17605/OSF.IO/WERK2 |
| Software and algorithms | | |
| SPM8 | Wellcome Trust Centre for Neuroimaging | https://www.fil.ion.ucl.ac.uk/spm/software/spm8/; RRID:SCR_007037 |
| SPM12 | Wellcome Trust Centre for Neuroimaging | https://www.fil.ion.ucl.ac.uk/spm/software/spm12/; RRID:SCR_007037 |
| fMRIPrep | Poldrack Lab | https://fmriprep.org/en/stable/; RRID:SCR_016216 |
| CANLab Core Tools | Wager Lab | https://github.com/canlab/CanlabCore/ |

## RESOURCE AVAILABILITY

### Lead contact

Further information and requests for resources should be directed to and will be fulfilled by the Lead Contact, Philip Kragel (pkragel@emory.edu).

### Materials availability

This study did not generate new unique reagents.

### Data and code availability

fMRI data are available at https://osf.io/werk2/. MATLAB code for analyses is available at: https://github.com/canlab.

## EXPERIMENTAL MODEL AND SUBJECT DETAILS

### 7T fMRI Study (Expt 1)

This study included 56 participants ($M_{age}$ = 26.46 years, SD = 5.87 years, 27 female). All recruited participants were between the ages of 18 and 40 years, were right-handed, had normal or corrected to normal vision, were not pregnant, were fluent English speakers, had no known neurological or psychiatric illnesses, and were recruited from the greater Boston area. Participants were excluded from the study if they were claustrophobic or had any metal implants that could cause harm during scanning. All participants provided written informed consent and study procedures were completed as approved by the Partners' Healthcare Institutional Review Board.

### 3T fMRI Study (Expt 2)

This study included 48 adult participants (mean ± SD age: 25.1 ± 7.1; 21 female, 27 male; 7 left-handed; 40 white and 8 non-white (1 Hispanic, 5 Asian, 1 Black, and 1 American Indian)). All participants were healthy, with normal or corrected to normal vision and normal hearing, and with no history of psychiatric, physiological or pain disorders and neurological conditions, no current pain symptoms, and no MRI contraindications. Eligibility was assessed with a general health questionnaire, a pain safety screening form, and an MRI safety screening form. Participants were recruited from the Boulder/Denver Metro Area. The institutional review board of the University of Colorado Boulder approved the study, and all participants provided written informed consent.

## METHOD DETAILS

### 7T fMRI Study (Expt 1)

#### *Experimental Paradigm*

Participants completed a probabilistic avoidance learning task during fMRI. In this task (Roy et al., 2014), participants learned to associate different visual cues (circles or triangles) with the aversiveness of stimuli (i.e., mechanical stimulation to the bed of the thumb (N = 31) or unpleasant visual images from the IAPS (Lang et al., 1997) (N = 25). Mechanical stimuli were delivered at non-painful

Please cite this article in press as: Kragel et al., A human colliculus-pulvinar-amygdala pathway encodes negative emotion, Neuron (2021), https://doi.org/10.1016/j.neuron.2021.06.001





(3 kg/cm$^2$) or painful (5 kg/cm$^2$) pressure levels. Visual stimuli were randomly selected from negative (normative valence = 2.93 ± 0.67 SD) and neutral (normative valence = 5.52 ± 0.57 SD) IAPS images. 24 trials of each type were presented in each of five runs (120 trials total). Participants we presented with both visual cues on every trial, followed by a decision phase, a jittered interstimulus-interval, and probabilistic reinforcement based on their decision. Reinforcement rates were predetermined based on a random walk between 20% and 80%, with outcomes determined randomly on every trial.

#### *MRI data acquisition and preprocessing*

Gradient-echo echo-planar imaging BOLD-fMRI was performed on a 7 tesla Siemens MRI scanner. Functional images were acquired using GRAPPA-EPI sequence: echo time = 28 ms, repetition time = 2.34 s, flip angle = 75°, number of slices = 123, slice orientation = transversal (axial), anterior to posterior phase encoding, voxel size = 1.1 mm isotropic, gap between slices = 0 mm, field of view = 205 × 205 mm$^2$, GRAPPA acceleration factor = 3; echo spacing = 0.82 ms, bandwidth = 1414 Hz per pixel, partial Fourier in the phase encode direction: 7/8. A custom-built 32-channel radiofrequency coil head array was used for reception. Radiofrequency transmission was provided by a detunable band-pass birdcage coil. Structural images were acquired using a T1-weighted EPI sequence: echo time = 22 ms, repetition time = 8.52 s, flip angle = 90°, number of slices = 126, slice orientation = transversal (axial), voxel size = 1.1 mm isotropic, gap between slices = 0 mm, field of view = 205 × 205 mm$^2$, GRAPPA acceleration factor = 3; echo spacing = 0.82 ms, bandwidth = 1414 Hz per pixel, partial Fourier in the phase encode direction: 6/8. This sequence was selected so that functional and structural data would have similar spatial distortions to facilitate coregistration and subsequent normalization of data.

Results included in this manuscript come from preprocessing performed using FMRIPREP (Esteban et al., 2019), a Nipype (Gorgolewski et al., 2011) based tool. Spatial normalization to the ICBM 152 Nonlinear Asymmetrical template version 2009c was performed through nonlinear registration with the antsRegistration tool of ANTs v2.1.0 (Avants et al., 2008; Gorgolewski et al., 2011), using brain-extracted versions of both T1w volume and template. Brain tissue segmentation of cerebrospinal fluid (CSF), white matter (WM) and gray matter (GM) was performed on the brain-extracted T1w using fast (Zhang et al., 2001) (FSL v5.0.9). Functional data was slice time corrected using 3dTshift from AFNI v16.2.07 (Cox, 1996) and motion corrected using mcflirt (FSL v5.0.9). This was followed by co-registration to the corresponding T1w using boundary-based registration (Greve and Fischl, 2009) with six degrees of freedom, using flirt (FSL). Motion correcting transformations, BOLD-to-T1w transformation and T1w-to-template (MNI) warp were concatenated and applied in a single step using ants ApplyTransforms (ANTs v2.1.0) using Lanczos interpolation.

### 3T fMRI Study (Expt. 2)

#### *Experimental Paradigm*

Participants received five different types of stimulation (mechanical pain, thermal pain, aversive auditory, aversive visual, and pleasant visual), each at four stimulus intensities. 24 stimuli of each type (6 per intensity level) were presented over six fMRI runs in random order, with different stimulus types intermixed within runs. Following stimulation on each trial, participants made behavioral ratings of their subjective experience. Participants were instructed to answer the question 'How much do you want to avoid this experience in the future?'. Ratings were made with a visual analog rating scale, with anchors 'Not at all' and 'Most' displayed at the ends of the scale.

#### *Stimuli*

Visual stimulation was administered on the MRI screen and included normed pictures from the International Affective Picture System (Lang et al., 1997; see Table S4). We created four 'stimulus intensity levels' by selecting seven images per intensity level in a two-step process: preliminary selection based on normed valence ratings (averaged across male and female raters) from the IAPS database; and final selection based on ratings by 10 lab members (5 male, 5 female) in response to the question ''How aversive is this image? 1-100.'' The chosen images included photographs of animals (7), bodily illness and injury (12), and industrial and human waste (9). Each picture was presented for 10 s.

Pressure stimulation was administered using an in-house pressure pain device. This MRI-safe device provides dynamically controlled pressure stimulation using LabView software (National Instruments, Austin, TX). Four pressure levels were applied to an applicator placed on the left thumbnail for 10 s each (Level 1: 4 kg/cm2, Level 2: 5 kg/cm2, Level 3: 6 kg/cm2; Level 4: 7 kg/cm2).

Thermal stimulation was administered using an ATS Pathway System (Medoc Ltd., Haifa, Israel) with a 16 × 16 mm Peltier-based contact thermode. Four stimulus intensity levels were delivered to the thenar eminence of the left hand (Level 1: 45°C, Level 2: 46°C, Level 3: 47°C, Level 4: 48°C). Each thermal stimulus lasted 10 s (1.5 s ramp-up, 1.5 s ramp-down, 7 s at target temperature).

Auditory stimulation was administered using MRI-compatible headphones. We used the sound of a knife scraping on a bottle, which is a reliable aversive auditory stimulus (Kumar et al., 2008). Four stimulus intensity levels were created by scaling the amplitude of a single audio file (Level 1: Lv4 −8 dB, Level 2: Lv4 −4 dB, Level 3: Lv4 −1 dB, Level 4: original amplitude). Auditory stimuli lasted 10 s each.

## QUANTIFICATION AND STATISTICAL ANALYSIS

### MRI data acquisition and preprocessing

Whole-brain fMRI data were acquired on a 3T Siemens MAGNETOM Prisma Fit MRI scanner at the Intermountain Neuroimaging Consortium facility at the University of Colorado, Boulder. Structural images were acquired using high-resolution T1 spoiled gradient recall images (SPGR) for anatomical localization and warping to standard MNI space. Functional images were acquired with a

multiband EPI sequence (TR = 460 ms, TE = 27.2 ms, field of view = 220 mm, multiband acceleration factor = 8, flip angle = 44°, 64 × 64 image matrix, 2.7 mm isotropic voxels, 56 interleaved slices, phase encoding posterior > anterior). Six runs of 7.17 mins duration (934 total measurements) were acquired. Stimulus presentation and behavioral data acquisition were controlled using the psychophysics toolbox (Brainard, 1997) for MATLAB (The MathWorks, Inc., Natick, MA).

fMRI data were preprocessed using an automated pipeline implemented by the Mind Research Network, Albuquerque, NM. The preprocessing steps included: distortion correction using FSL's top-up tool (https://fsl.fmrib.ox.ac.uk/fsl/fslwiki/), realignment (affine alignment of first EPI volume (reference image) to T1, followed by affine alignment of all EPI volumes to the reference image and estimation of the motion parameter file (sepi_vr_motion.1D, AFNI; https://afni.nimh.nih.gov/), spatial normalization of the T1 image (T1 normalization to MNI space (nonlinear transform), normalization of EPI image to MNI space (3dNWarpApply, AFNI; https://afni.nimh.nih.gov/), interpolation to 2 mm isotropic voxels (for better alignment with templates in standard MNI152 space to facilitate prospective testing) and smoothing with a 6 mm FWHM kernel (SPM 8; https://www.fil.ion.ucl.ac.uk/spm/software/spm8/).

Prior to first level analysis, we removed the first four volumes to allow for image intensity stabilization. We also identified image-wise outliers by computing both the mean and the standard deviation (across voxels) of intensity values for each image for all slices to remove intermittent gradient and severe motion-related artifacts.

### fMRI data analysis

Data were analyzed using SPM12 (https://www.fil.ion.ucl.ac.uk/spm) and custom MATLAB code available from the authors' website (https://github.com/canlab/CanlabCore). First-level general linear model (GLM) analyses were conducted in SPM12. The six runs were concatenated for each subject. Boxcar regressors, convolved with the canonical hemodynamic response function, were constructed to model periods for the 10 s stimulation and 4-7 s rating periods. The fixation cross was used as an implicit baseline. A high-pass filter of 0.008 Hz was applied. Nuisance variables included: regressors coding for each run (intercept for each run); linear drift across time within each run; the six estimated head movement parameters (x, y, z, roll, pitch, and yaw), their mean-centered squares, their derivatives, and squared derivative for each run (total 24 columns); and motion outliers (spikes) identified in the previous step. A single-trial model was used to uniquely estimate the response to every stimulus to assess functional connectivity.

### Multivariate pathway identification

Functional connectivity between the amygdala, pulvinar, and superior colliculus was estimated using Partial Least-squares (PLS) regression, which identifies latent multivariate patterns that maximize the covariance between two blocks of data (Wold et al., 2001). This approach for estimating the covariance between two putatively connected neural populations, which we call multivariate pathway identification (MPathI; see Figure S1), is an extension of multivariate methods that have recently been proposed to estimate functional connectivity. These methods identify how patterns of activity in one region relate to patterns in another region, and produce estimates of connectivity based on multivariate dependence (Anzellotti and Coutanche, 2018; Anzellotti et al., 2017; Basti et al., 2020; Coutanche and Thompson-Schill, 2013). We note that MPathI is not an estimation procedure, but a framework for studying connectivity. A family of multivariate methods could be used for estimation, including PLS, canonical correlation, or other multivariate regression approaches. We were motivated to develop a new approach (that utilizes PLS for estimation) because it is uniquely aligned with our goal of identifying functionally connected neural populations, based on the following considerations: 1) it estimates the spatial distribution of underlying neural populations, unlike representational similarity and distance-based methods (Basti et al., 2020; Geerligs et al., 2016) that more flexibly search for statistical dependence; 2) it is constrained to estimate connectivity using robust, reproducible signals by reducing model complexity, which is a problem for related methods such as canonical correlation analysis (Le Floch et al., 2012; Smith et al., 2015); 3) it does not aim to explain all of the variance in each region, as is done when estimating multivariate pattern dependence (Anzellotti et al., 2017), but characterizes signals that covary between regions; and 4) it is a linear estimation procedure, which increases interpretability. These properties allow PLS to constrain the flexibility of connectivity models and allow them to be validated against what is known about neural pathways in animal models.

We used SIMPLS as developed by de Jong (1993) to identify latent scores ($\boldsymbol{T}$ and $\boldsymbol{U}$) and loadings ($\boldsymbol{P}$ and $\boldsymbol{Q}$) that maximize the covariance between the centered variables $\boldsymbol{X_0}$ and $\boldsymbol{Y_0}$:

1) Compute the cross-product $\boldsymbol{S} = \boldsymbol{X_0}'\boldsymbol{Y_0}$
2) Compute the singular value decomposition of $\boldsymbol{S}$
3) Get weights: $\boldsymbol{r}$ = first left singular vector
4) Compute $\boldsymbol{X}$ scores: $\boldsymbol{t} = \boldsymbol{X_0 r}$
5) Compute $\boldsymbol{X}$ loadings: $\boldsymbol{p} = \boldsymbol{X_0}'\boldsymbol{t}/(\boldsymbol{t}'\boldsymbol{t})$
6) Compute $\boldsymbol{Y}$ loadings: $\boldsymbol{c} = \boldsymbol{Y_0}'\boldsymbol{t}$
7) Compute $\boldsymbol{Y}$ scores: $\boldsymbol{u} = \boldsymbol{Y_0 c}$

With two additional steps to finding regression coefficients that map activity in one block onto the latent scores of the second block (where $\boldsymbol{A}^+$ is the Moore-Penrose pseudoinverse of matrix $\boldsymbol{A}$):

Please cite this article in press as: Kragel et al., A human colliculus-pulvinar-amygdala pathway encodes negative emotion, Neuron (2021), https://doi.org/10.1016/j.neuron.2021.06.001





8) Compute $\boldsymbol{Z} = \boldsymbol{X}^{+}\boldsymbol{u}$
9) Compute $\boldsymbol{V} = \boldsymbol{Y}^{+}\boldsymbol{t}$

These coefficients can be used to make predictions on out-of-sample data, e.g., in cross-validation or prospective tests.

In Expt. 1 preprocessed BOLD time-series data were used as inputs for pathway models, and separate connectivity models were fit for each subject. To assess generalizability across subjects and stimulus types in Expt. 2, PLS models were fit using single-trial estimates of BOLD responses (Gazzaley et al., 2004) to aversive thermal, mechanical, auditory, and visual stimuli, in addition to a set of pleasant visual stimuli which were used as a control. These data were concatenated across subjects (5,760 total trials, 120 per subject). For the pulvinar–amygdala model, the predictor (***X***) block of variables included all 208 voxels in an anatomically defined mask of the pulvinar and the outcome (***Y***) block included all 240 voxels in the amygdala (Amunts et al., 2005; see Figure S2 for depiction of anatomical ROIs). The superior colliculus–pulvinar model included responses in all voxels in a hand-drawn mask of the superior colliculus (41 voxels) as the predictor block and pulvinar responses as the outcome.

This set of regions is particularly well suited to pattern-based estimation of connectivity because each region has a fine-grained structure with unique profiles of connectivity. The superior colliculus is composed of multiple layers, with superficial layers predominantly receiving visual inputs and deeper layers integrating multisensory information and coordinating orienting behavior (Tardif et al., 2005; May, 2006). Both the pulvinar and amygdala are composed of multiple subnuclei, which have distinct profiles of subcortical and cortical connectivity (Arcaro et al., 2018; Barron et al., 2015; Chang and Yu, 2018; Gattass et al., 2017; Guedj and Vuilleumier, 2020; Hrybouski et al., 2016). Pattern-based analysis of fMRI activity can capture information coded in neural substrates exceeding the resolution of single voxels, as it samples population activity distributed across voxels, acting as a complex spatiotemporal filter (Kriegeskorte et al., 2010). Because anatomy varies across individuals, and population activity may be blurred due to hemodynamic filtering, our masks were not set at a conservative threshold (i.e., including any voxels that could plausibly be identified as a specific region), anticipating that MPathI would identify that the key regions within each region that covary with one another.

Each PLS model was specified to identify a pair of latent variables that maximally covary with one another. The patterns of activity in each region that predict latent activity in the other region were estimated using least-norm regression. Inference on these patterns was made using bootstrap resampling (5,000 iterations, treating each subject as a block and randomly sampling subjects with replacement) with normal approximation for inference. Leave-one-subject-out cross-validation was performed to estimate the strength of functional connections (using the Pearson correlation between actual and estimated brain activity as the measure of interest). Inference on cross-validated estimates of functional connectivity were performed using block permutation tests (1,000 iterations) in which the order of trials in the outcome block were scrambled independently for each subject.

### Sensitivity analysis

The sensitivity of the brain pathway models was evaluated by comparing the correlation of multivariate model estimates (e.g., correlations of latent activity common to the pulvinar and amygdala) to that of the mean signals in the same regions. We tested for differences in correlation coefficients by converting them to z-scores using the Fisher transform and performing a t test of differences between pathway models and the mean signal.

### Control analyses

To evaluate whether the MPathI approach can classify pathways with known monosynaptic connections from those that do not, we constructed a set of connectivity models between thalamic nuclei and primary sensory cortex. These included one pathway model from the lateral geniculate nucleus (LGN, from the Morel Atlas; Krauth et al., 2010) to primary visual cortex (V1, from a multimodal parcellation of the cortex; Glasser et al., 2016) and another from the medial geniculate nucleus (MGN, from the Morel atlas) to primary auditory cortex (A1 from Glasser et al., 2016). We also estimated the connectivity of "crossed" connections (i.e., from LGN to $V1_{MGN}$ and from MGN to $A1_{LGN}$) based on the activity of latent patterns in each region. We used these four estimates to classify established pathways (i.e., LGN-V1 and MGN-V1) from indirect pathways. Inference on the strength of individual pathways was performed using one-sample t tests on Fisher transformed correlation coefficients.

### Mediation analysis

A pathway from the pulvinar to the amygdala is thought to mediate emotional responses by rapidly conveying information about the environment from the superior colliculus. To test this hypothesis, we performed a series of mediation analyses identifying three statistical paths to characterize the effects of the activity of different brain pathways on normative aversiveness: 1) path *a* characterizes the effect of the $SC_{pulv}$ pathway on $Pulv_{Amy}$ activity; 2) path *b* reflects the relationship between $Pulv_{Amy}$ activity and normative ratings of aversiveness, and 3) path *ab* reflects pulvinar activity formally mediating the link between superior colliculus activity and normative aversiveness, reducing the strength of the direct path *c'* between these variables. We performed mediation analyses to assess relationships between brain activity and normative differences in aversiveness because they are more closely linked to the function of the colliculus-pulvinar-amygdala pathway, i.e., to rapidly evaluate threatening sensory cues. Compared to trial-by-trial self-reports, normative ratings are more directly linked to the nature of stimuli, less influenced by decision variables, including social and contextual factors, and are less influenced by individual differences in introspection and self-report, which are likely mediated by other brain systems including insular cortices and the default network (Chang et al., 2013; Gu et al., 2013; Kleckner et al., 2017). We tested the

generalizability of this pathway using mediation analysis both for negative images and sounds and assessed specificity by performing the same mediation analysis on responses to painful stimuli (thermal and mechanical) and positive images. We additionally evaluated whether this pathway is sensitive to moment-to-moment differences in affective experience, as opposed to the normative intensity of negative valence, by performing the same analysis with on-line ratings.

### Cortical and thalamic pathway estimation

To evaluate whether activity in the colliculus-pulvinar-amygdala pathway is distinct from cortical activity during emotion processing, we estimated pathway models between the amygdala and parcels of the Glasser atlas combined across hemispheres (Glasser et al., 2016) and anatomically defined regions of the thalamus (Krauth et al., 2010). This produced a series of beta estimates for 197 regions. These estimates were concatenated into a single 5,760 (trial) by 197 (region) matrix that was subjected to principal component analysis (PCA). Bootstrap resampling (5,000 samples, with block resampling of subjects to keep all images from a subject together) was performed to estimate the standard errors of coefficients from PCA.

The activity of each component was estimated by multiplying the PCA coefficients and activation estimates for all pathways. The association between each component and subjective aversiveness (and pleasantness) ratings was estimated using the procedures described in the sensitivity analysis section.