Ha Ha! Versus Aha! A Direct Comparison of Humor to Nonhumorous Insight for Determining the Neural Correlates of Mirth

Ori Amir1, Irving Biederman1,2, Zhuangjun Wang2 and Xiaokun Xu1

1Department of Psychology, 2Neuroscience Program, University of Southern California, Los Angeles, CA 90089-1061, USA

Address correspondence to Ori Amir, Department of Psychology, University of Southern California, SGM 501, 3620 South McClintock Ave, Los Angeles, CA 90089-1061, USA. Email: oamir@usc.edu

While humor typically involves a surprising discovery, not all discoveries are perceived as humorous or lead to a feeling of mirth. Is there a difference in the neural signature of humorous versus nonhumorous discovery? Subjects viewed drawings that were uninterpretable until a caption was presented that provided either: 1) a nonhumorous interpretation (or insight) of an object from an unusual or partial view (UV) or 2) a humorous interpretation (HU) of the image achieved by linking remote and unexpected concepts. fMRI activation elicited by the UV captions was a subset of that elicited by the humorous HU captions, with only the latter showing activity in the temporal poles and temporo-occipital junction (linking remote concepts), and medial prefrontal cortex (unexpected reward). Mirth may be a consequence of the linking of remote ideas producing high—and unexpected—activation in association and classical reward areas. We suggest that this process is mediated by opioid activity as part of a system rewarding attention to novel information.

Keywords: Cortical µ-opioid gradient, fMRI, humor, medial prefrontal cortex, temporal pole

Introduction

Humor involves a discovery of an unexpected interpretation or perspective similar to what occurs when we experience insight. However, not all insight experiences are funny. Previous imaging studies of humor have contrasted humorous stimuli with the same stimuli without the humorous element (e.g., Mobb et al. 2003; Moran et al. 2004; Bartolo et al. 2006; Kohn et al. 2011), and/or compared different kinds of humor (e.g., Goel and Dolan 2001; Watson et al. 2007; Samson et al. 2008, 2009; Franklin and Adams 2011). Unfortunately, a comparison among subtypes of humor necessarily ignores the regions involved in all humor, and a comparison of humorous stimuli to the same stimuli with the humorous element removed may be too broad, as controls lack the element of discovery. The present investigation directly compares humorous versus nonhumorous cognitive discovery.

More recently studies have compared humor with more subtle controls, for example, humorous clips versus nonhumorous stimuli that inspire a positive/enjoyable feeling (Neely et al. 2012; Vrticka et al. 2013), humorous resolvable incongruities versus nonresolvable incongruities (Samson et al. 2008, 2009; Bekinschtein et al. 2011; Chan, Chou, Chen, Liang 2012; Chan, Chou, Chen, Yeh et al. 2012). Yet in all of those studies, an element of discovery was arguably present in the humorous conditions but not in the controls, or at least, was not systematically controlled for.

The imaging studies of humor cited above generally reported activation in portions of the temporal lobes often near the junctions with the parietal or occipital lobes and/or the temporal poles (TPs) and some classical reward regions. Additional regions were also reported, but not consistently across the different studies. The inconsistency might be the result of a lack of control for the element of discovery, as well as additional confounds, common in comedy, which enhance the reward, but are not necessary for humor, such as schadenfreude, superiority, or sexual titillation (Hurley et al. 2011). In the present investigation, we use humorous stimuli that were not designed to induce these emotions. In some of the prior experiments, as noted by Bartolo et al. (2006), humor relied heavily on attribution of intentions, eliciting activity in the regions associated with theory of mind. Our stimuli did not rely heavily on such attributions.

Several imaging studies have explored the neural basis of insight (e.g., Bowden et al. 2005; Aziz-Zadeh et al. 2009). Owing to the large variation in the nature of the stimuli, the tasks, and the controls (ranging from anagrams to physical problem solving involving object interactions), no consistent pattern of activation emerged across the various studies (for a review, see Dietrich and Kanso 2010).

To study the neural correlates of mirth (enjoying humor) while controlling the element of discovery, subjects in the present investigation viewed 2 kinds of simple line drawings, that were, by themselves, uninterpretable with respect to their possible referents (Fig. 1). A subsequently presented caption would be provided that elicited, for one condition, an insight-like (but nonhumorous) interpretation of an object depicted from a partial or unusual view (UV) (Fig. 1A,D) by providing the referent for that object (UV condition). Responses to UV stimuli were compared with responses to the second kind of drawing/caption, with a humorous interpretation (HU) (Fig. 1C,D), based on a surprising linking of remote associations (HU condition). Each HU and UV stimulus had a control caption that merely provided a physical description of the drawing, adding no referential information. For each drawing, half the subjects viewed the original caption and half the control caption. Critically, both UV and HU conditions (unlike their controls) result in a discovery, that is, the referent or a new interpretation, of the previously uninterpretable line drawing.

Materials and Methods

Subjects viewed simple line drawings (Fig. 1) whose referent was inaccessible in the absence of a caption. To compare the brain’s response with humor and insight, we compared the activation elicited from humorous “doodles” (Price 1955, 1976, 2000), such as the ones shown in Figure 1C,D, in which the humor is based on the unexpected linking of remote concepts, to drawings of objects in partial or unusual views (Nishimoto et al. 2010; e.g., Fig. 1A,D) defining an insight condition where the description provides the referent but does not elicit a humorous response. There is some disagreement in the literature whether the label “insight” should be applied to instances in which the
solution is provided, as is the case in our experiment, or only when subjects reach the solution on their own. However, in the present experiment providing the solution for the insight condition renders that condition more comparable with the humorous droolles condition, as the HU is similarly provided (rather than asking the subjects to come up with it themselves).

Participants
Fifteen adults, 7 females, all right handed (except for 1 male), age 19–31 (mean = 22.6). All were students at the University of Southern California, except for one graduate from another university. We obtained informed consent from all subjects, and they were compensated for their participation. The study was approved by the University Park Institutional Review Board at the University of Southern California.

Stimuli
The stimuli were 200 line drawings along with both interpretive and physical descriptions. One hundred images whose original captions provided the HU condition were scanned from 3 books by Price (1955, 1976, 2000), whose drawings with their captions he termed "Droodles." We used only Droodles that were readily understood and did not require knowledge of out of date conventions or objects. The remaining stimuli, 100 (insight) images whose original captions referred to objects depicted from a partial or UV interpretation were taken from Nishimoto et al. (2010). Out of their database of 196 exemplars created for experimental purposes, 100 were selected based on interpretability.

Interpretive descriptions were generally taken as written by the original authors. Some minor changes were made to the descriptions of UVs to make them more accessible to American subjects. Physical descriptions were created which described the drawing physically without conveying its interpretation and which, on average, approximated the length of the drawing’s interpretive description with respect to the number of words.

Subjects viewed both types of drawings, HUs and UVs, followed by either the referential or the control descriptions for a total of 4 conditions. Each drawing’s 2 descriptions were counterbalanced between subjects so that half the subjects saw the control description and half saw the referential description. A particular drawing was viewed in only a single condition by a given subject.

<table>
<thead>
<tr>
<th>Drawing:</th>
<th>Unusual View Interpretation (Insight) Drawing:</th>
<th>Humorous Interpretation Drawing:</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>Thick T-junctions</td>
<td>Two smaller horizontal ellipses in a larger vertical ellipse.</td>
</tr>
<tr>
<td>B</td>
<td>Rectangles with square tabs</td>
<td>A plethora of dots surrounded concentrically around a single dot</td>
</tr>
<tr>
<td>C</td>
<td>Trumpet Valves</td>
<td>Close-up of a pig looking at book titles in a library.</td>
</tr>
<tr>
<td>D</td>
<td>Fluorescent light bulbs.</td>
<td>Germs avoiding a friend who caught antibiotics.</td>
</tr>
</tbody>
</table>

Figure 1. Examples of the 4 conditions. Credits: (A and B): Adopted from Nishimoto et al. (2010). (C and D): Droodles excerpts from "Droodles - the Classic Collection" © 2000 by Tallfellow press, Los Angeles. Used by Permission. All rights reserved. (C): Droodle orientation and caption differs from the original Droodle.

Procedure
Each 15-s trial began with a 2-s display of a drawing (subtending a visual angle of ~8°) in the center of the screen. The drawing was shown alone to pique the subject’s curiosity (Fig. 2). Then for 5 s, either a physical or interpretive description was displayed below the drawing followed by 3 s of just the drawing itself, allowing for the interpretation to be fully appreciated. Trials were separated by a 2-s blank interval. Presentation sequences were programmed with Psychophysics Toolbox (Brainard 1997; Pelli 1997) running on MATLAB (The MathWorks, Natick, MA, USA). Subjects were instructed to rate each drawing for funniness on a scale of 1–4 with 4 buttons in the scanner: 1 indicated they did not understand the description and/or its relationship to the drawing; 2–4 indicated the degree of funniness with 2 being “not funny,” 3 “a little funny,” and 4 “funny.” A pretest had determined that “funny” was more effective than “very funny” in encouraging subjects to use the full scale. Responses were collected from the onset of the description to the end of the trial. Note that only one-fourth of the trials were designed to be funny (HU type), while

Figure 2. The temporal arrangement of a trial in the fMRI experiment. Droodles excerpts from “Droodles - the Classic Collection” © 2000 by Tallfellow press, Los Angeles. Used by Permission. All rights reserved.
another one-fourth of the trials were interpretive but of the UVs type, and one-half the trials were the control physical descriptions.

In one session, each subject participated in an anatomical scan and 4 experimental runs of 50 drawings, each lasting 10 min.

**Data Acquisition**

All fMRI images were scanned at USC's Dana and David Dornsife Cognitive Neuroscience Imaging Center on a Siemens Trio 3T scanner with a standard 16-channel head coil. Each subject ran in a high-resolution T2-weighted structural scan using MPRAGE sequence. (Repetition time (TR) = 1100 ms, 192 sagittal slices, 256 x 256 matrix size, 1 x 1 x 1 mm voxels).

Functional images were acquired using an echo-planar imaging (EPI) pulse sequence with the parameters: TR = 2000 ms, TE = 30 ms, flip angle = 62°, 256 x 256 matrix size, in plane resolution 3 x 3, 3 mm thick slices, 32 axial slices covering as much of the brain as possible, always including the TPs, but occasionally missing the superior rim of the primary motor and somatosensory cortices.

Functional data were corrected for motion artifacts. We then conducted a whole-brain, random-effects fixed effect analysis, because of the large window of time selected for the event-related activity for both conditions.

We found that the regions showing differential activity for both UVs versus controls (Table 2). Among those regions were the following: left frontal regions largely overlapping with Broca's language area, a small portion of the posterior inferior frontal gyrus on the right, previously suggested to involve holding multiple alternative meanings in language processing (Mashal et al. 2005), and an area largely overlapping with bilateral lateral occipital complex (LOC)—a region selective for visual images of object shape (Grill-Spector et al. 1999; Hayworth and Biederman 2006). The parahippocampal gyrus (PHG), a region which activates more strongly to visual scenes compared with single objects (Epstein and Kanwisher 1998), and responds more strongly to visual objects and scenes that produce rich contextual associations (Bar et al. 2008), was activated more for both interpretive conditions.

### Results

**Behavioral Results**

Subjects rated a drawing as “1” (I do not understand) on only 1.3% of the trials, and those trials were removed from further analysis. Mean ratings for the 4 conditions are shown in Table 1. The HU drawings were rated as significantly funnier than the UV interpretation drawings, t(14) = 12.39, P < 0.001, and the physical description control of the HU condition (HUcont), t(14) = 12.42, P < 0.001. UV drawings, which were not meant to be humorous, still received a slightly, but significantly, higher mean ratings than their UVcontrol, t(14) = 2.93, P = 0.011.

Mean response times for responding to the HU drawings was 3.64 s (SD = 0.86), which included the time to read the description. This time was significantly longer than that for UV, M = 2.92, SD = 1.15, t(14) = 5.04, P < 0.001. The RTs for the control conditions HUcontrol, M = 3.08, SD = 1.12, and UVcontrol = 2.51, SD = 1.18, were both significantly shorter than those of their respective interpretive conditions, respectively t(14) = 4.53 and t(14) = 7.01, both P's < 0.001. The 400- to 600-ms RT difference between HU and UV may reflect the longer time required to get the referent for the HU stimuli and, perhaps, also deciding its funniness. Owing to the large window of time selected for the imaging data analysis (TRs 4–9, TR = 2 s), we likely captured all of the event-related activity for both conditions.

<table>
<thead>
<tr>
<th>Condition</th>
<th>Mean funniness rating</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>HU (humorous)</td>
<td>3.351</td>
<td>0.236</td>
</tr>
<tr>
<td>UV (Insight)</td>
<td>2.220</td>
<td>0.246</td>
</tr>
<tr>
<td>HU control</td>
<td>2.090</td>
<td>0.159</td>
</tr>
<tr>
<td>UV control</td>
<td>2.035</td>
<td>0.068</td>
</tr>
</tbody>
</table>

### Table 1

Mean funniness ratings and SDs for the 4 conditions

Funniness ratings were 2—“not funny,” 3—“a little funny,” and 4—“funny.” Ratings of 1—“I don’t understand” are excluded.

**Humorous Interpretation Drawings and Unusual View Interpretation Drawings versus Controls**

The HU and UV drawings were contrasted with their physical description controls. Figure 3 displays these contrasts on an inflated cortical map (see also Fig. 4, for additional regions not apparent on the inflated maps). We used the P < 0.005 threshold for all statistical analysis. We found that the regions exhibiting greater activation for the UVs versus their controls were a subset of those activated for the HUs versus their controls.

A conjunction analysis revealed the regions activated by both UV and HU conditions compared with their respective controls (Table 2). Among those regions were the following: left frontal regions largely overlapping with Broca's language area, a small portion of the posterior inferior frontal gyrus on the right, previously suggested to involve holding multiple alternative meanings in language processing (Mashal et al. 2005), and an area largely overlapping with bilateral lateral occipital complex (LOC)—a region selective for visual images of object shape (Grill-Spector et al. 1999; Hayworth and Biederman 2006). The parahippocampal gyrus (PHG), a region which activates more strongly to visual scenes compared with single objects (Epstein and Kanwisher 1998), and responds more strongly to visual objects and scenes that produce rich contextual associations (Bar et al. 2008), was activated more for both interpretive conditions.

Increased activation for both interpretive conditions was also observed in the bilateral ventral striatum, and the left amygdala (IAMG) (while the right amygdala was significantly activated for both HU–HUcontrol and UV–UVcontrol, no right amygdala activation survived either the conjunction of these 2 contrasts or their subtraction (see Supplementary Tables 1 and 2)).
both of which are part of the dopaminergic reward network (Lee et al. 2004), and both were reported in previous fMRI studies of humor (e.g., IAMG specifically was reported in Mobbs et al. (2003), Bartolo et al. (2006), Watson et al. (2007), although some studies found bilateral amygdala activity, e.g., Moran et al. (2004) and still others reported no amygdala activity). Our findings suggest that activation in these 2 reward regions occurs in response to nonhumorous discovery as well (note, however, that a small region within the left striatum was significantly more activated for HU than UV [see Fig. 4, top left image, and Table 3]). Activity in both the PHG and the striatum has been found to correlate positively with preference for visual scenes (Yue et al. 2007).

Finally, a region extending from the dorsal portions of the anterior cingulated cortex (dACC) to the supplementary motor area (SMA) was activated for both interpretive conditions. Mobbs et al. (2003) reported activation in this region for humorous stimuli, and they suggested it to be involved in the motor planning of laughter. However, that region’s elevated activity for the nonhumorous UV condition suggests an alternative cognitive function: the reassessment of the display based on the new interpretation (an executive function often attributed to ACC, specifically in studies of insight; e.g., Aziz-Zadeh et al. (2009)).

**Regions Only Activated for Humorous Interpretations**

A subtraction of activation for UV drawings from HU drawings (after each condition’s activation was subtracted from its respective control: [HU − HU_cont] − [UV − UV_cont]), revealed 4...
regions of activation (see Fig. 4 top and Table 3): 1) The medial prefrontal cortex (mPFC), an area found to be involved in reward learning that responds to unexpected reward contingencies (Rolls 1996), and also has been found to correlate with funniness ratings of stimuli in previous studies (e.g., Goel and Dolan 2001), 2) the TPJ, and 3) the temporoparietal junctions (TOJs) extending to 4) temporoparietal junctions (TPJs). Mobbs et al. (2003) found activation in both TOJ and TP as well, but only on the left side, and suggested TPJ activity may relate to lexical retrieval or semantic knowledge processing, and the TOJ does semantic processing and may be detecting incongruity/surprise. TPJ activity, particularly in the right hemisphere has been linked to mentalizing; however, the left TPJ is more responsive to other similar cognitive tasks not requiring modeling of other minds, that is, perspective taking and comparing interpretations (Perner et al. 2006). Regardless of their exact function in the context of humor, these are association regions involved in high-level semantic processing (Meyer and Damasio 2009; Man et al. 2012). While there is some inconsistency in the humor literature against which we have compared our results, as discussed in the Introduction, this may be the result of the variable nature of controls (which often either included some funny elements or were lacking the element of discovery). Unlike previous studies, we compared the humorous HU condition with a nonhumorous UV condition which retained the element of discovery, allowing regions responsive to humor to be distinguished from those responsive to insight or reinterpretation.

### Eliminating Potential Confounds

The HU Drawings compared with those for the UVs were more complex visually—as there were, on average, more pixels in the drawings, and their verbal interpretation tended to be longer. Consequently, the comparisons of HU and UV interpretive conditions were always made after subtracting from the interpretive conditions their respective descriptive controls. The descriptive controls for the HUs and UVs were such that the images were identical and the captions were of the same length (see Materials and Methods). To assess whether any of the differences between the HUs and UVs minus their respective controls were the result of differences in figure complexity or descriptive length/content, we examined the contrast between the 2 physical controls: HUcont – UVcont. This contrast only yielded significant activation in early visual areas up to and including LOC (but not PHG) and in the left hemisphere frontal language areas (using the same P < 0.005, uncorrected, as in all of the analysis we report). Note that none of these regions were the ones differentiating HU from UV activations.

Another concern was that laughter, in the humorous HU condition, would result in head movements that might artificially increase the BOLD signal in some of the regions we reported. While head movements were small to begin with (always <2 mm, and typically <0.5 mm), and they were corrected in the preprocessing (see Materials and Methods), we conducted an additional test to confirm that head movements did not explain our findings: we added the 6 motion correction parameters (3D translations and rotations) to the design matrix of the GLM in BrainVoyager, and obtained virtually identical results with exclusion of the motion correction parameters in the design matrix. All results presented here were computed with the design matrix including the motion correction parameters.

### Ratings of Humor

We contrasted the activation, within the interpretive HU condition, for drawings that were rated by each subject as 4 (funny), minus those that were rated 3 (a little funny) (Fig. 5) to assess whether a “dose–response” effect would be evident and, if so, whether this effect would also involve those regions differentially activated by the humorous minus the nonhumorous insight conditions. These comparisons included fewer trials (because they constituted only the trials from the HU drawings, and only those rated either 3 or 4). Nevertheless, under the same threshold (P < 0.005), robust activation was evident in the same regions that differentiated HU from UV conditions in the previous analyses: bilateral TPJs, the temporoparietal junction, and bilateral medial prefrontal-frontal cortex. However, in this analysis, the region of differential activation-labeled mPFC extended more dorsally within the mPFC, and it was more prominent on the left side. The striatum

---

**Table 2**

<table>
<thead>
<tr>
<th>ROI</th>
<th>No. of voxels</th>
<th>X</th>
<th>Y</th>
<th>Z</th>
<th>Mean ± t</th>
<th>Mean ± P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>rLOCC</td>
<td>8571</td>
<td>38</td>
<td>−53</td>
<td>−26</td>
<td>4.090</td>
<td>0.002</td>
</tr>
<tr>
<td>rFrontal</td>
<td>5042</td>
<td>−47</td>
<td>20</td>
<td>1</td>
<td>4.071</td>
<td>0.002</td>
</tr>
<tr>
<td>rCer</td>
<td>4588</td>
<td>30</td>
<td>−61</td>
<td>−33</td>
<td>4.131</td>
<td>0.002</td>
</tr>
<tr>
<td>ISTR</td>
<td>4035</td>
<td>−15</td>
<td>−7</td>
<td>11</td>
<td>3.876</td>
<td>0.002</td>
</tr>
<tr>
<td>rLOC</td>
<td>3789</td>
<td>−41</td>
<td>−49</td>
<td>−22</td>
<td>3.772</td>
<td>0.003</td>
</tr>
<tr>
<td>rADCC</td>
<td>3736</td>
<td>−5</td>
<td>12</td>
<td>56</td>
<td>3.825</td>
<td>0.002</td>
</tr>
<tr>
<td>rSTR</td>
<td>2036</td>
<td>11</td>
<td>−3</td>
<td>11</td>
<td>3.781</td>
<td>0.002</td>
</tr>
<tr>
<td>rAMG</td>
<td>1316</td>
<td>−30</td>
<td>−6</td>
<td>−7</td>
<td>3.897</td>
<td>0.002</td>
</tr>
<tr>
<td>rIPG</td>
<td>981</td>
<td>−32</td>
<td>−34</td>
<td>−22</td>
<td>3.915</td>
<td>0.002</td>
</tr>
<tr>
<td>rPHG</td>
<td>468</td>
<td>28</td>
<td>−40</td>
<td>−24</td>
<td>3.968</td>
<td>0.002</td>
</tr>
<tr>
<td>rFrontal</td>
<td>342</td>
<td>55</td>
<td>19</td>
<td>24</td>
<td>3.748</td>
<td>0.003</td>
</tr>
</tbody>
</table>

Note: ROIs are defined using a threshold of P = 0.005, uncorrected. ROI sizes in voxels as well as Talairach coordinates, average t and P values are provided.

**Table 3**

<table>
<thead>
<tr>
<th>ROI</th>
<th>No. of voxels</th>
<th>X</th>
<th>Y</th>
<th>Z</th>
<th>Mean ± t</th>
<th>Mean ± P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>rTP</td>
<td>4989</td>
<td>38</td>
<td>14</td>
<td>−25</td>
<td>4.113</td>
<td>0.002</td>
</tr>
<tr>
<td>rTPJ</td>
<td>3570</td>
<td>42</td>
<td>−62</td>
<td>11</td>
<td>4.437</td>
<td>0.001</td>
</tr>
<tr>
<td>lmPFC</td>
<td>3543</td>
<td>−7</td>
<td>48</td>
<td>40</td>
<td>4.118</td>
<td>0.002</td>
</tr>
<tr>
<td>rTP</td>
<td>3139</td>
<td>−50</td>
<td>3</td>
<td>−23</td>
<td>3.991</td>
<td>0.002</td>
</tr>
<tr>
<td>rTOJ</td>
<td>2545</td>
<td>57</td>
<td>−56</td>
<td>−2</td>
<td>3.876</td>
<td>0.002</td>
</tr>
<tr>
<td>rOFC</td>
<td>1133</td>
<td>28</td>
<td>11</td>
<td>−18</td>
<td>3.864</td>
<td>0.002</td>
</tr>
<tr>
<td>rTPJ</td>
<td>1013</td>
<td>−57</td>
<td>−61</td>
<td>6</td>
<td>3.823</td>
<td>0.002</td>
</tr>
<tr>
<td>Neg_rParietal</td>
<td>640</td>
<td>17</td>
<td>47</td>
<td>54</td>
<td>3.792</td>
<td>0.002</td>
</tr>
<tr>
<td>lTOJ</td>
<td>528</td>
<td>−61</td>
<td>−49</td>
<td>−3</td>
<td>3.947</td>
<td>0.002</td>
</tr>
<tr>
<td>rCEr</td>
<td>515</td>
<td>20</td>
<td>−38</td>
<td>−33</td>
<td>4.032</td>
<td>0.002</td>
</tr>
<tr>
<td>rHIp</td>
<td>430</td>
<td>25</td>
<td>−32</td>
<td>−16</td>
<td>3.929</td>
<td>0.002</td>
</tr>
<tr>
<td>lSTR</td>
<td>375</td>
<td>−7</td>
<td>−11</td>
<td>10</td>
<td>3.785</td>
<td>0.003</td>
</tr>
<tr>
<td>Neg_rFrontal</td>
<td>362</td>
<td>28</td>
<td>29</td>
<td>38</td>
<td>−3.838</td>
<td>0.002</td>
</tr>
<tr>
<td>Neg_lSTG</td>
<td>311</td>
<td>41</td>
<td>−34</td>
<td>8</td>
<td>−3.702</td>
<td>0.003</td>
</tr>
</tbody>
</table>

Note: ROIs are defined using a threshold of P = 0.005, uncorrected. ROI sizes in voxels as well as Talairach coordinates, average t and P values are provided.

---
and parahippocampal cortices were also activated bilaterally, as well as most of the left cingulate cortex, precuneus, and lingual gyri on both sides. This overlap between regions activated for higher funniness ratings and the regions specified by comparing the HU minus the UV conditions suggests that the latter difference indeed amounts to a manipulation of humor.

Previous imaging studies of humor that had both a nonhumorous control and obtained funniness ratings did not report that all regions showing higher activation for the humorous condition, relative to the nonhumorous control, also showed a dose response (e.g., Goel and Dolan 2001; Watson et al. 2007). The fact we did find a dose response suggests that we have tighter controls over this complex behavior.

Discussion

The “Humor” Network

Both the humorous HU drawings, and the insight-like UV drawings had an element of discovery and stimulus reinter-pretation. However, a network of 4 regions, uniquely and bilaterally active for the humorous stimuli was revealed by subtracting activation for the UV from the HU drawings: the TPs, temporo-occipital junction (TOJ), extending to the temporoparietal junction (TPJ), and the mPFC. The role of this network in humor processing is further supported by its greater activation in the separate analysis comparing, within the HU condition, those drawings that were rated funniest with those rated as less funny.

What Makes Humor Humor?

In an often cited (almost by default) attempt to characterize the cognitive process underlying humor appreciation, Suls (1972) suggested that humor is a form of problem solving in which there are 2 stages: 1) a perception of incongruity between the “punch line” and what was expected, and 2) a resolution of the incongruity. However, that also characterizes nonhumorous insight (Ruch and Hehl 1998). What, then, distinguishes the 2 experiences? Navon (1988) suggested that jokes are characterized by a situation/action that is appropriate from one perspective typically assumed by the protagonist of the joke, but virtually inappropriate or absurd in reality. Similarly, Ruch and Hehl (1998) suggested a third stage to Suls’ theory, in which there is a realization that the resolution does not make sense (is only an “as if” resolution). Hurley et al. (2011) suggested that all humor is based on realizing an inappropriate interpretation/ perception was reached because of an erroneous assumption/heuristic. McGraw and Warren (2010) more generally described the conditions for humor as the perception of a benign violation. What is common to the 4 theories above is that some absurd, erroneous, or inappropriate perception/behavior is entertained resulting in violated expectations. Indeed, the funniest drawings elicited activity in the anterior cingulate cortex, which has been suggested to encode errors/conflicts (Botvinick et al. 2004; Brown and Braver 2007; Shenhav et al. 2013). These accounts also imply that, in order for these conditions to be met, humor needs to be based on remote associations—associations not readily elicited by the set up—simply because otherwise it might be impossible to mislead or surprise the listener. The linking of the remote associations (e.g., in Fig. 1C, a pig looking at book titles in the library) likely occurs in association cortex where its novelty, the violation of expectations, and/or the rejection of the (apparently appropriate) link as absurd (e.g., because pigs do not read), necessarily results in heightened activity in association regions if the elements are to be actively maintained so that they can be linked by the punch line (or caption, in our study). Once we hear a joke, its repetition is no longer funny as a conceptual structure has been formed that has already integrated the previous remote associations.

Accordingly, we find that these humorous stimuli induce activation in association regions, namely bilateral TOJ, TPJ, and TP, where information converges, and where remote associations may be integrated.

Could the same explanation apply to any other arbitrary constellation of regions? We believe it could not, since this constellation of regions is unique in satisfying 2 requirements: 1) it consists of higher order association areas where remote associations converge to give rise to semantic interpretations; 2) the regions have been reported in previous independent investigations of humor. (Although those regions have been reported in previous investigations, they generally were reported as part of a larger constellation of regions, and our work reduces the set of relevant regions, by controlling for the element of discovery.)

The Element of Discovery

Unlike previous studies in which only the humorous stimuli retained an element of discovery, both the humorous HU and nonhumorous (insight) UV conditions in our study retained an element of discovery, that is, the drawings were uninterpretable until their descriptions revealed what they depicted. Regions that are activated for the UV condition (minus its descriptive controls) appear to be largely a subset of those activated by the HU condition (minus its controls). Our findings (see The Humor Network), therefore, suggest reconsideration of at least 2 interpretations of prior studies.
Watson et al. (2007) found that, late in the processing of their jokes, sight gags led to an increased activation in visual areas, and verbal gags led to increased activation in language areas. They interpreted the findings as support for Suls’ (1972) incongruity-resolution theory of humor, suggesting the late visual or linguistic activity arises from the resolution of the visual or linguistic incongruity, respectively. Goel and Dolan (2001) used similar results to reach the same conclusion. However, we found a late increase in activation of LOC (a visual object recognition region) for both HU and UV conditions compared with their controls (Fig. 6). In other words, we found the same activation pattern of “resolution” for both our humorous HU condition and the nonhumorous UV condition, results that challenge Suls’ characterization of the conditions for humor.

Second, as discussed in the Results and Discussion, since the SMA was activated for both HU and UV, it seems less likely its only function in humor processing is laughter production (as was suggested by Mobbs et al. 2003).

As noted, humor typically has an element of absurdity or benign violations (e.g., Navon 1988). In visual humor, typical depictions might include talking animals or a person in an incongruous pose. That is, even without the caption, such depictions look funny—and would have imposed a serious confound in our effort to distinguish humor from discovery. Fortunately, the HU droodles depicted no such absurdities. Instead, like the control UV stimuli, they appeared “abstract” with no real-world referent. It was the caption, not the drawing, that conveyed the absurdity of the referent that thus rendered the composite funny. We could thus assess the impact of the captions alone, unconfounded with the images.

Why is Humor Pleasurable?
Many theories have suggested evolutionary advantages for enjoying humor, such as, a motivation for error debugging (Hurley et al. 2011) and a number of previous studies reported activation in some classical reward areas (e.g., Mobbs et al. 2003). We are aware of only one, however, that suggested a neural mechanism linking the cognitive processing of humor and the pleasurable feeling of mirth (e.g., Biederman and Vessel 2006).

There is near consensus in the field that “surprise” is an element of humor, and that the surprise is positive (or at least benign). Often the surprise induces a feeling of superiority, Schadenfreude, relief, or sexual titillations, but these factors are not necessary (Hurley et al. 2011). None of these were aspects of our humorous stimuli. The positive surprises induced by our stimuli seem to be the “cleverness” of the interpretation. Indeed, in a separate survey, 20 student subjects rated the same HUs used in the fMRI study, and the survey revealed a high correlation of Funniness ratings with ratings of both Cleverness ($r=0.823$, $P<0.001$) and Surprise ($r=0.711$, $P<0.001$) but not Complexity ($r=-0.122$, $P=0.188$). The link between cleverness and pleasure may be an aspect of a general motivational system, which renders the consumption of novel and richly interpretable information as pleasurable.

Lewis et al. (1981) discovered a gradient of $\mu$-opioid receptors in the ventral cortical visual pathway, sparse in early sensory areas, but gradually increasing in density through higher sensory cortices, with the greatest density in association cortex (e.g., PHG, TPs). Zadina et al. (1997) reported the presence of $\mu$-opioid receptors in prefrontal cortex. Biederman and Vessel (2006) hypothesized that neural activation in association cortex would be greater in response to richly interpretable, novel, and surprising experiences, resulting in greater opioid release, possibly triggering further activation of classical reward regions, producing pleasure. They proposed that such activation of cortical opioids might be the neural basis for the human motivation for seeking novel, richly interpretable information (for additional evidence, see Yue et al. 2007; Amir et al. 2011).

The TPs and TOJs association areas were significantly activated for HUs but not UVs. Also responding exclusively to HUs, the mPFC, which is implicated in reward learning and
responds most strongly to unexpected rewards (or their absence; Rolls 1996; Lee et al. 2004). This characterization fits well with the unexpected and rewarding nature of punch lines.

In conclusion, the mirthful feeling accompanying humor appreciation may come about as a result of unexpected greater activation in cortical association regions rich in opioid receptors (e.g., bilateral TOJ and TP), triggering, in response to this unexpected reward, activation in classical reward/reward-learning regions such as the mPFC, IAMG, and striatum (with the latter 2 regions also activated for nonhumorous discovery).

**Summary**

A handful of regions differentiate the discovery inherent in humor from nonhumorous discovery. The same regions were also more responsive to funnier humorous stimuli, thus exhibiting a dose–response effect. We hypothesized that the surprisingly high activation in association regions (e.g., TP and TOJ) leads to the high activation in reward regions (e.g., mPFC) in line with previous studies (e.g., Biederman and Vessel 2006) that linked greater associative activity with perceptual pleasure.

**Supplementary Material**

Supplementary material can be found at: http://www.cercor.oxfordjournals.org/

**Funding**

This work was supported by National Science Foundation BCS 04-20794, 05-31177, and 06-17699 to I.B.

**Notes**

We thank Mark Lescoart for sharing his Matlab code, and Jianchung Zhuang for his assistance with the MRI scanning. Conflict of Interest: None declared.

**References**


