Social Regulation of Affective Experience of Humor

Vinod Goel\textsuperscript{1,2} and Raymond J. Dolan\textsuperscript{1}

Abstract

The element of surprise, a necessary condition for the experience of humor, often derives from the fact that the alternative interpretation/resolution offered by the punch line of a joke is physically or socially forbidden. Children’s humor typifies violation of physical norms, whereas adult humor typically pushes the boundaries of social norms. Excess norm violation, to the point of offending, can attenuate the experience of humor/mirth. To examine the neural basis of regulation of affective experience of humor by social norms, we scanned 16 normative subjects while they viewed a series of cartoons that varied in funniness and social acceptability. Behavioral results indicated two separate groups of subjects, those who found the cartoons less offensive and those who found them more offensive. In the group that found the jokes more offensive, there was a negative correlation between funniness and social inappropriateness. In this group, the corresponding humor by social inappropriateness interaction during functional magnetic resonance imaging revealed enhanced activation in the right hippocampus along with relative deactivation in the ventral medial prefrontal cortex (VMPFC). By contrast, the funniness by social appropriateness interaction resulted in activation in the VMPFC and relative deactivation in the right hippocampus. These results suggest that the regulation of humor by social norms involves reciprocal response patterns between VMPFC and hippocampus regions implicated in contextual regulation of behavior and memory, respectively.

INTRODUCTION

A necessary (but not sufficient) condition for humor is a juxtaposition of mental sets. This involves setting up a situation that points to a particular interpretation, and then introducing a twist or “punch line” that offers an unexpected, but meaningful, juxtaposition leading to an alternative interpretation. When successful, this juxtaposition can lead to a rewarding experience of mirth, well-being, and laughter. Goel and Dolan (2001) identified prefrontal—temporal systems as being involved in juxtaposition of mental sets (i.e., in getting the joke) and a ventral medial prefrontal cortex (VMPFC) system as being a critical component for the affective experience of successful jokes (i.e., finding the joke funny), and other labs (Osaka & Osaka, 2005; Moran, Wig, Adams, Janata, & Kelley, 2004; Mobbs, Greicius, Abdel-Azim, Menon, & Reiss, 2003) have also identified basal ganglia structures, including the nucleus accumbens in the affective experience in humor.

The surprise element in humor often derives from the fact that the alternative interpretation offered by the punch line is either physically impossible or socially forbidden. Indeed, children’s humor is largely predicated on the violation of physical laws. For example, in the popular “Road Runner” cartoon series, the coyote routinely falls 1000 m down a cliff followed by an anvil and a piano that always land on top of him. He always picks himself up, pushes down the large “bump” on his head, and walks away to fall another day. To entertain this scenario, we need to suspend our beliefs about what is physically possible and impossible. In contrast to children’s humor, adult humor tends to rely heavily on pushing the boundaries of social norms, often involving sexual scenarios (e.g., Q: How do you attract the most desirable women with just 7 words? A: “Do you take VISA or American Express?”). However, although pushing of social norms is necessary, excessive norm violation, to the point of offending, can dampen the experience of humor/mirth (e.g., Q: What’s yellow and green, stinks, and lies on the side of the road? A: A dead Girl Scout.). Thus, social norms serve to regulate the affective experience of humor.

To investigate the neural basis of the social regulation of the affective experience of humor, we scanned 16 right-handed normal subjects using event-related functional magnetic resonance imaging (fMRI), while they viewed cartoons varying in levels of social acceptability. Based on previous fMRI and patient studies, we expected the experience of mirth to be associated with activation in the VMPFC (Goel & Dolan, 2001; Rolls, 2000; Schultz, Tremblay, & Hollerman, 2000; Anderson, Bechara, Damasio, Tranel, & Damasio, 1999), and social norm violation to activate the orbital frontal cortex (OFC) and the amygdala (Moll, de Oliveira-Souza, Bramati, & Grafman, 2002; Anderson & Phelps, 2001; Garavan, Pendergrass, Ross, Stein, & Risinger, 2001;
Anderson, Bechara, Damasio, Tranel, & Damasio, 2000; Davidson, Jackson, & Kalin, 2000; Lemerise & Arsenio, 2000; Adolphs, Tranel, & Damasio, 1998; Morris et al., 1998).

**METHODS**

**Subjects**

Sixteen right-handed normal subjects (4 men), with no history of psychiatric and neurological disorders and a mean age of 26.94 years (SD = 10.01 years) and mean education level of 17.12 years (SD = 2.76 years), volunteered to participate in the study. All subjects gave informed consent, and the study was approved by the Joint National Hospital for Neurology and Neurosurgery/Institute of Neurology Ethics Committee.

**Stimuli Presentation**

Stimuli were cartoons gathered from the Internet. Cartoons typically consist of a picture and a verbal caption. Usually, the picture provides a context and the linguistic caption delivers the punch line (i.e., you first look at the picture and then read the caption). Reversing the order often diminishes the effect. However, we were interested in delivering a pictorial punch line (to compare and contrast the results with a previous study that used only linguistic material; Goel & Dolan, 2001). Based on prescan behavioral pilot studies (with subjects of similar age and education as the fMRI subjects), we selected 92 cartoons (from 800 + items) that encompassed the spectrum of funny and socially acceptable, funny and socially unacceptable, not funny and socially acceptable, and not funny and socially unacceptable, and worked well when presented in the reverse order (linguistic caption followed by pictorial punch line). Social acceptability was violated by the presence of sexual material. Examples of more and less socially acceptable trials appear in Figure 1. Stimuli were resized to a uniform size (30 in.²) and presented in random order.

Each 10,980-msec trial began with a trial marker, followed by the text caption at 1000 msec. At 4000 msec, the picture punch line appeared. At this point, subjects were requested to answer the following two questions: (i) How funny do you find the cartoon on a scale of 1 to 7? (ii) How socially appropriate do you find the cartoon on a scale of 1 to 7? These questions required subjective judgments that were recorded by keypresses. Subjects’ responses were used to rate stimuli on a subject by subject, trial by trial basis. Examples and presentation details of stimuli are provided in Figure 1A and B.

**fMRI Scanning Technique**

A 1.5-T Siemens Sonata system (Siemens, Erlangen, Germany) was used to acquire T1 anatomical volume images (1 1 1.5 mm voxels) and 36 T2*-weighted echo-planar images (64 64 3 mm pixels, TE = 40 msec) sensitive to blood-oxygenation-level-dependent (BOLD) contrast. Echo-planar images (1.8 mm thick) were acquired axially every 3 mm and positioned to cover the whole brain. Data were recorded during a single acquisition period. A total of 318 volume images were acquired in one session with a repetition time (TR) of 3.24 sec/volume. The first six volumes were discarded to allow for T1 equilibration effects.

Trials from all conditions were randomly presented in a single-event design. There were 92 event presentations during the session, each having a duration of 11 sec (for a session length of 16.87 min). The scanner was synchronized with the presentation of all trials in the session.

**Data Analysis**

Data were analyzed using Statistical Parametric Mapping (SPM2) (Friston et al., 1995). All volumes were spatially realigned to the first volume (head movement was < 3 mm in all cases) and temporally realigned to the AC–PC slice, to account for different sampling times of different slices. A mean image created from the realigned volumes was coregistered with the structural T1 volume and the structural volumes were spatially normalized to the Montreal Neurological Institute brain template (Evans et al., 1993) using nonlinear basis functions (Ashburner & Friston, 1999). The derived spatial transformation was then applied to the realigned T2 volumes, which were finally spatially smoothed with a 12-mm FWHM isotropic Gaussian kernel (in order to make comparisons across subjects and to permit application of random field theory for corrected statistical inference; Worsley & Friston, 1995). The resulting time series across each voxel were high-pass filtered with a cutoff of 32 sec, using cosine functions to remove section-specific low-frequency drifts in the BOLD signal. Global means were normalized by proportional scaling to a grand mean of 100, and the time series were temporally smoothed with a canonical hemodynamic response function to swamp small temporal autocorrelations with a known filter.

All events were modeled in the design matrix, but the event of interest was the neural activity at the presentation of the cartoon punch line. The BOLD signal time locked to the presentation of the cartoon punch line was modeled as a hemodynamic response function. The other events (i.e., the presentation of the “setup” line and the decisions and motor responses corresponding to the judgments on degree of social acceptability and funniness) were modeled as events of no interest (with hemodynamic response functions). Condition effects at each voxel were estimated according to the general linear model and regionally specific effects compared using linear contrasts. Each contrast produced a statistical parametric map of the t statistic for each voxel,
which was subsequently transformed to be unit normal Z-distribution. In the absence of a priori hypotheses, we report only activations surviving a voxel-level intensity threshold of \( p < .05 \) (corrected for multiple comparisons in a random effect model using False Discovery Rate) (Genovese, Lazar, & Nichols, 2002). For specific regions for which we have a priori hypotheses, we use the threshold of \( p < .001 \) (uncorrected).

RESULTS

Each subject rated the jokes for funniness and social acceptability on a scale of 1 to 7. The mean rating for funniness was 3.88 (SD = 0.59) and 4.06 (SD = 0.88) for social acceptability. Overall, there was no significant correlation between funniness and social acceptability \( r(14) = .302, \text{ns} \), suggesting independence of the two indices. However, a closer examination of the behavioral data revealed two distinct subgroups of subjects: Those who were more offended by the material and those who were less offended. Dividing the subjects into these two groups of eight showed a significant difference between the high \( (M = 4.26, \text{SD} = 0.55) \) and low \( (M = 3.32, \text{SD} = 0.36) \) groups in social appropriateness \( t(14) = 4.06, p < .01 \). The group that found the jokes socially appropriate also found them funnier \( (M = 4.17, \text{SD} = 0.51) \) than the group that found them inappropriate \( (M = 3.20, \text{SD} = 0.60); t(14) = 3.41, p < .01 \). In the group that found the material more socially acceptable, there was no correlation between funniness and social acceptability ratings \( r = .20, \text{ns} \). In the group that found the material more offensive \( (M = 3.32, \text{SD} = 0.36) \), there was a significant correlation between funniness and
social acceptability ($r = .32, p < .01$). What this suggests is that, in those subjects who considered the stimuli, as a whole, socially appropriate, the issue of funniness was independent of the issue of social inappropriateness. However, in those subjects whose ratings indicated more sensitivity to the socially inappropriate trials, to the point of being offended, increasing levels of norm violation attenuated the effect of funniness.

Given these behavioral results, we carried out parametric analysis of the fMRI data to examine the overall effect of humor and social norm violation, and then divided the subjects into two groups, those who were more offended ($n = 8$) and those who were less offended ($n = 8$), for further analysis. Across the entire group, we isolated neural responses covarying with funniness by performing a parametric analysis using subjects’ subjective ratings of funniness and found that the medial frontal cortex (BA 10) ($2, 56, 8; Z = 3.71$), including the ventral medial orbital cortex (BA 11) ($0, 42, 22; Z = 2.92$); the left insula ($42, 24, 18; Z = 3.79$); the basal ganglia, including the nucleus accumbens ($14, 3, 0; Z = 3.07$); the right superior temporal gyrus (BA 22) ($60, 50, 20; Z = 3.87$); and the right cerebellum showed increasing activity with increasing funniness ratings (see Figure 2). Activity in the left inferior OFC (BA 11) ($10, 56, 16; Z = 3.91$ and $34, 32, 18; Z = 3.26$), right inferior temporal gyrus (BA 37) ($52, 70, 2; Z = 4.17$), right amygdala ($36, 2, 16; Z = 3.32$), and left cuneus (BA 18) ($10, 102, 8; Z = 4.49$) increased with increasing ratings of social inappropriateness (see Figure 3).

Within the more offended group of subjects, a Funniness Social norm interaction corresponding to the behavioral correlation was evident. The Funny Social violation interaction (calculated by multiplying social norm violation ratings by funniness ratings and using this measure as a covariate in a parametric analysis) was

Figure 2. A statistical parametric map (SPM) rendered into standard stereotactic space and superimposed on to sagittal (A) and coronal (C, D) sections of a magnetic resonance image, which is, in itself, in standard space. (A) Activation in the MPFC (BA 10) ($2, 56, 8; Z = 3.71$) including the VMMPFC (BA 11) ($0, 42, 22$) covaried with funniness ratings. (B) Parameter estimate graphs (generated by grouping funniness ratings into high, medium, and low categories) show increasing activation in the VMMPFC ($0, 42, 22$) in response to increasing funniness of cartoons.

Figure 3. A statistical parametric map (SPM) rendered into standard stereotactic space and superimposed on to sagittal (A) and coronal (C, D) sections of a magnetic resonance image, which is, in itself, in standard space. (A) Activation in the left orbital cortex (BA 11) ($0, 42, 22$) and (B) ($34, 32, 18$) ($2, 56, 16; Z = 3.91$) covaried with increasing social norm violation. (C, D) Parameter estimate graphs (generated by grouping social inappropriateness ratings into high, medium, and low categories) show increasing activation in the left orbital frontal cortex (C) ($34, 32, 18$) and (D) ($10, 56, 16$) in response to increasing social norm violation.
expressed as a relative enhanced activation in the right hippocampus (32, 14, 12; Z = 4.44; 44, 8, 26; Z = 4.63) and relative deactivation in the VMPFC (6, 42, 8; Z = 3.65). By contrast, the reverse Funny Social appropriateness interaction (calculated by multiplying the inverted social norm violation ratings by funniness ratings and using this measure as a covariate in a parametric analysis) resulted in activation in the VMPFC and right insula and relative deactivation in the right hippocampus (see Figure 4). Thus, although the orbital PFC is activated by socially inappropriate material, an excessive violation of social norms results in activation of the right hippocampus and relative deactivation in the medial PFC in association with the affective regulation of humor. There was no interaction in the unoffended group.

DISCUSSION

The modulation of response in the medial PFC and VMPFC, left insula, and right superior temporal gyrus (and nucleus accumbens at a reduced threshold), with the subjects’ subjective rating of funniness of jokes, is consistent with previously reported fMRI results. Several studies have implicated the right temporal lobe in processing and maintaining alternative, less probable meanings, and in integrative processes requiring global coherence (Goel & Dolan, 2001; Coney & Evans, 2000; Federner & Kutas, 1999; St George, Kutas, Martinez, & Sereno, 1999; Faust & Chiarello, 1998). These are the types of processes necessary for semantic juxtapositions of meaning required to "get" the punch line. The insula has previously been reported as activated during the experience of mirth with verbal jokes (Goel & Dolan, 2001), cartoons (Mobbs et al., 2003), and film (Moran et al., 2004).

The finding that regions of the VMPFC and nucleus accumbens are activated in the appreciation of funny jokes is consistent with previous findings (Mobbs et al., 2003; Goel & Dolan, 2001) and these regions’ involvement in mesolimbic reward systems (Rolls, 2000; Schultz et al., 2000; Anderson et al., 1999). The fact that the Goel and Dolan (2001) study used verbal jokes and the present study uses cartoons further reinforces the role of the VMPFC as a modality-independent region interested in reward processing.

There are only a few lesion studies of humor appreciation (Docking, Murdoch, & Jordan, 2000; Shammi &
Stuss, 1999; Shields, 1991; Braun, Lussier, Baribeau, & Ethier, 1989; Gardner, Ling, Flamm, & Silverman, 1975). They widely report right hemisphere involvement, but the nature of the tests administered, variability in location, and etiology of human lesions preclude these studies from being more specific about component processes and functional anatomy (Shammi & Stuss, 1999; Braun et al., 1989; Wapner, Hamby, & Gardner, 1981). Our results suggest that what these studies may be capturing is the right temporal lobe's involvement in integrative semantic processes requiring global coherence, necessary for juxtaposition of mental sets (St George et al., 1999; Brownell, Simpson, Bihrlie, Potter, & Gardner, 1990; Brownell, Potter, Michelow, & Gardner, 1984). A study of 10 frontal lobe patients reported that 6 of 10 patients—all with lesions to the VMPFC—had diminished humor responses (Shammi & Stuss, 1999), a finding consistent with our results.

In contrast to the experience to funniness, the experience of social norm violation activated the left orbital frontal cortex, the right posterior temporal gyrus, the right amygdala, and the left cuneus. The OFC has been implicated in social processing in both imaging and patient studies (Moll et al., 2002; Stone, Cosmides, Tooby, Kroll, & Knight, 2002; Anderson, Bechara, et al., 2000; Anderson, Damasio, Tranel, & Damasio, 2000; Davidson et al., 2000; Lemerise & Arsenio, 2000), whereas the role of the amygdala in social judgment (Moll et al., 2002; Adolphs et al., 1998) and the perception and processing of various forms of emotional informational cues (Anderson & Phelps, 2001; Garavan et al., 2001; Davidson et al., 2000; Morris et al., 1998) is widely recognized. Furthermore, the amygdala, OFC, and posterior temporal gyrus are regions that are constitutive of the so-called social brain, and their engagement with norm violation is consistent with their putative role in social processing.

The novel contribution of the present study is the finding of an interaction between social norm violation and funniness in the more offended subgroup of subjects. In this “more offended” subgroup, there was a negative correlation between funniness and social norm violation ratings, such that subjects found jokes that minimally violated social norms funnier than those that violated social norms more extensively. The neural response of this subgroup to higher “funnier” ratings in the more socially acceptable material was expressed in an increased activation in the VMPFC and right insula, a result similar to that of the overall group. However, the more extensive norm violation trials crossed some acceptability threshold in these subjects, attenuating their experience of mirth. The neural response to increasing social norm violation was expressed as increasing activation in the left OFC, as with the overall group. However, the funny by social norm violation interaction, indicating an attenuated “funniness” ratings in the face of increasing norm violation, resulted in enhanced activation in the right hippocampus and (a relative) decrease in activation in the M/VMPFC. These results suggest that although the VMPFC is activated by humorous situations involving acceptable social norm violation, an excessive violation of social norms triggers shock (correlated with activation in the right hippocampus), which in turn attenuates the experience of humor/mirth (manifested as relative deactivation in the VMPFC). This pattern suggests a potential neural substrate for the regulation of the affective experience of humor by social context. In keeping with this interpretation, it is interesting to note that there are strong connections between the hippocampus and medial sectors of the OFC, which could provide an anatomical substrate for reciprocal interactions between these structures (Cavada, Company, Tejedor, Cruz-Rizzolo, & Reinoso-Suarez, 2000). However, as a caveat, it is worth remembering that the current study does not provide a direct measure of functional or effective connectivity between these regions. This remains an objective of future research.

Reprint requests should be sent to Vinod Goel, Department of Psychology, York University, Toronto, Ontario, Canada M3J 1P3, or via e-mail: vgoel@yorku.ca.

Notes

1. We ran regressions to examine the linear and curvilinear relationships between rated funniness and social appropriateness. The results demonstrated that the linear ($b = .47, t = 2.04, p < .05$) and curvilinear ($b = .56, t = 2.07, p < .05$) relationships were both present, suggesting that this particular breakdown is capturing a meaningful factor.

2. There were no differences in the “funniness” and “appropriateness” ratings between the male and female subjects. In particular, the four male subjects were evenly distributed between the two (offended, underfunded) groups (with mean ratings of 2.67, 3.6, 3.64, and 5.12).

3. We are using the term “offended” loosely. Our probe was about social appropriateness. The trials deemed socially inappropriate may or may not have offended all subjects.

REFERENCES


