



The elephant in the room: attention to salient scene features increases with comedic expertise

Ori Amir¹ · Konrad J. Utterback¹ · Justin Lee¹ · Kevin S. Lee² · Suehyun Kwon¹ · Dave M. Carroll² · Alexandra Papoutsaki²

Received: 13 August 2021 / Accepted: 25 January 2022

© Marta Olivetti Belardinelli and Springer-Verlag GmbH Germany, part of Springer Nature 2022

Abstract

What differentiates the joke writing strategy employed by professional comedians from non-comedians? Previous MRI work found that professional comedians relied to a greater extent on “bottom-up processes,” i.e., associations driven by the prompt stimuli themselves, while controls relied more on prefrontal lobe directed, “top-down” processes. In the present work, professional improv comedians and controls generated humorous captions to cartoons while their eye movements were tracked. Participants’ visual fixation patterns were compared to predictions of the saliency model (Harel et al. in *Adv Neural Inf Process Syst* 19:545–552, 2007)—a computer model for identifying the most salient locations in an image based on visual features. Captions generated by the participants were rated for funniness by independent raters. Relative to controls, professional comedians’ gaze was driven to a greater extent by the cartoons’ salient visual features. For all participants, captions’ funniness positively correlated with visual attention to salient cartoon features. Results suggest that comedic expertise is associated with increased reliance on bottom-up, stimulus-driven creativity, and that a bottom-up strategy results, on average, in funnier captions whether employed by comedians or controls. The cognitive processes underlying successful comedic creativity appear to adhere to the old comedians’ adage “pay attention to the elephant in the room.”

Keywords Humor · Expertise · Saliency · Eye-tracking · Bottom-up processing

Introduction

What characterizes strategies employed by professional artists when generating their art? The neural basis for creativity has been the subject of much recent research, using functional magnetic resonance imaging (fMRI) to monitor brain activation of artists, such as music improvisers (e.g., Limb and Braun 2008; Villarreal et al. 2013), writers (e.g., Howard-Jones et al. 2005; Chen et al. 2019; Erhard et al. 2014), and painters (e.g., Aziz-Zadeh et al. 2013; Schlegel et al. 2015; Rominger et al. 2020). While the specific brain regions activated differed across artistic domains, the medial

prefrontal cortex (mPFC) was reported as a common region partaking during nearly all forms of creative efforts (Dietrich and Kanso 2010; Chen et al. 2020).

Amir and Biederman (2016) proposed that humor creativity may serve as an ideal case study for creativity in general, as jokes take a relatively short time to improvise (at least by professional improv comedians), their quality (i.e., funniness) can be readily judged, and expertise can be easily coded (i.e., by comparing professional comedians to amateurs and controls). The study found that while both the mPFC and semantic regions in the temporal lobes bilaterally (TMP) were activated during the improvisation of comedic captions to cartoons, only TMP activity was positively correlated with the captions’ funniness. While professional comedians relative to amateurs and controls showed increased TMP activation during comedic improvisation, mPFC activation was negatively correlated with comedic expertise (Amir and Biederman 2016; similar reduction but of dorsolateral prefrontal cortex activity with increased musical expertise was reported by Pinho et al. 2014). The authors suggested the role of the mPFC is to direct the

Handling Editor: Cees van Leeuwen (KU Leuven); Reviewers: Andrey Nikolaev (Lund University).

✉ Ori Amir
ori.amir@pomona.edu

¹ Psychological Science, Pomona College, Claremont, CA, USA

² Computer Science, Pomona College, Claremont, CA, USA

creative process in a task appropriate manner—a top-down process that is needed to a lesser extent for professionals, whose bottom-up processes are adapted to produce funny ideas more autonomously.

The terms “bottom-up” and “top-down” have long been used by cognitive scientists to describe the process by which the brain determines which aspects of a stimulus to attend to and in what manner to process them (Kinchla and Wolfe 1979). A bottom-up process is a stimulus-driven process, meaning the features of the stimulus itself determine to which aspects of the stimulus attention would be directed (the most *salient*¹ image features, by definition) and the nature of the processing of the stimulus. A top-down process is a goal-directed one, in which attention is actively directed to features which may be useful to a particular task and the nature of the processing of the stimulus is informed by the task’s goal. An example of a top-down process may involve visually searching one’s red-sweater wearing son in a crowd by actively looking for red features and attending to them, while a bottom-up process may involve being distracted by and directing attention to a worker wearing a bright yellow vest in the background (Theeuwes 2010). The neural basis of a bottom-up process is believed to be a feedforward process in which information propagates from primary sensory regions (e.g., visual cortex) toward higher level association regions and finally prefrontal cortex. In the process, features to be attended to are selected for further processing. Those features may subsequently be interpreted and recognized, and semantic associations may be generated, all in a feedforward manner (Hummel and Biederman 1992). A top-down process, on the other hand, involves feedback neural connections from areas such as the prefrontal cortex which actively directs attention, processing, and the generation of semantic associations in a manner which might best serve a goal (e.g., focus on red, to achieve the goal of finding one’s son; Itti 2000). What Amir and Biederman’s (2016) brain imaging study suggests is that while comedy writing involves both bottom-up (indexed by the degree of temporal activation) and top-down (indexed by mPFC activation) processing, in professional comedians, the degree of top-down processing is reduced in favor of bottom-up processing.

Humor theorists have yet to reach consensus on the necessary conditions for humor and on the conditions, which merely enhance it (Martin and Ford 2018; McGraw and Warren 2010; Amir 2016; Amir et al. 2013). But most humor theorists would agree that a common feature of jokes, possibly a necessary one, is an element of incongruity—often a contradiction between the expectations set by the setup vs.

the punchline. An incongruity that is commonly followed by a resolution, i.e., a revised understanding of the setup that is more consistent with the punchline (Suls 1972; Ruch et al. 1993). That process of reinterpretation is often reported to elicit an internal sense of “surprise,” a characteristic often associated most readily with humor by the “layperson” (Ruch 2001). However, it is important to note that the relationship between the incongruity-resolution process and the emotional reaction of surprise is mostly based on introspection, and, in fact, humor may be experienced in the absence of an unexpected element or a sense of surprise (e.g., in the case of a class of jokes that are still funny the second time they are heard; Hurley et al. 2011). Hurley et al. (2011) proposed that jokes always operate by violating expectations, *on some level*, but the expectations being violated need not be those of the audience (e.g., in a case of a practical joke, the victims’ expectations may be violated while the audience is fully expecting the punchline). Raskin (2012) proposed that for an expectation to be violated and an incongruity to be resolved a joke often makes use of two scripts. In the joke “how do you stop an elephant from charging? you take away its credit card” the setup elicits a primary script: an elephant *running* toward a person. The punchline elicits a secondary script of an elephant charging *money*. The primary script is the sense most readers of the joke would associate with the word “charging” in that context. The secondary script is unexpected by most and may elicit a sense of surprise (Raskin 2012). The process of writing a punchline for a setup or other prompt may involve considering how most individuals would initially interpret the prompt and then considering some alternative interpretations or more remote but relevant associations for the prompt—such exercise may be a fundamental aspect of comedic creativity (Shahaf et al. 2015). Indeed, Houston and Mednick (1963) proposed a fundamental aspect of creativity in general is the meaningful linking of *remote* associations.

In apparent contradiction to the emphasis on the incongruous or unexpected, humor is commonly believed to be funny “cause it’s true” and when it addresses the “elephant in the room”. Indeed, improv comedians are trained to “listen and react” by aiming their focus on the unfolding “scene” and “getting out of their head” which may be translated to a suggestion to rely more on bottom-up/stimulus-driven (and less on top-down) processes which are likely to result in jokes more responsive and relevant to the scene. In 2015, Shahaf et al. attempted to characterize the funniest captions of the thousands submitted to the New Yorker’s cartoon captioning competition. They found that captions which make use of associations to the theme of the cartoon that reside in the middle ground of accessibility—namely those associations that are not the most obvious, yet not obscure or irrelevant—result in funnier captions (Shahaf et al. 2015). An effective comedy writing process which may achieve both

¹ A saliency of a feature is often defined, in such context, to indicate the degree to which the feature attracts attention in the absence of a goal-oriented task (e.g., during free viewing; see Itti et al. 1998).

ends of a punchline that is prompt² relevant *and* unexpected may be accomplished by more robust bottom-up processing. A bottom-up (i.e., stimulus-driven) process should be more effective as it is likely to generate prompt *relevant* associations. A *robust* bottom-up process should be more effective as it is likely to generate *more* and *more remote* associations of which some may satisfy the conditions of being both prompt relevant and unexpected.

Given our understanding of the nature of humor and comedic creativity briefly summarized above, in what ways would we expect expertise in comedy writing to affect the manner of scanning a cartoon during the process of generating humorous captions to it? If the aim of the comedy writer is to misdirect the listener in order to generate a sense of incongruity or surprise, one may be tempted to hypothesize that seasoned comedians may look away from the most obvious/prominent elements of the image in order to think of non-obvious captions that may be more likely to be incongruous with the audience's initial understanding of the image. On the other hand, if there is a comedic imperative to address the most prominent features of the situation (a.k.a., “the elephant in the room”), one may hypothesize that professional comedians would spend *greater* time gazing at the salient features of the cartoon. The former strategy would require a top-down intervention, while the latter favors a more pronounced bottom-up process consistent with previous brain imaging results (Amir and Biederman 2016). As discussed above, a *robust* bottom-up strategy may actually achieve both ends in the following manner: an increased focus on the salient features may result in an overall increased number of relevant associations generated—of which some would satisfy the conditions of being both remote yet stimulus relevant associations—such associations have been suggested to be a better basis for creativity in general (Houston and Mednick 1963) and comedy in particular (Shahaf et al. 2015). If a bottom-up strategy is indeed more effective, we would expect professional comedians to gaze more time at the salient features of the cartoons during the process of contemplating funny captions. We would also expect that independent of comedic expertise, funnier captions would emerge, on average, following trials in which participants gazed more at the salient cartoon features.

How can the saliency of various cartoon features be quantified? In 1998, Itti et al. proposed the Saliency Model, a computational model for predicting visual fixations within an image based on the low-level visual features. According to the model, the features most likely to attract visual fixations are those that differ from other features within their region of the image (e.g., a vertical edge among horizontal

edges, a dark patch within a bright environment). The model predictive performance was further improved by Harel et al. (2007), for example, by predicting with increased likelihood fixations within the center of the image—due to increased visual acuity in foveal (central) region of the retina (Parkhurst et al. 2002; Hirsch and Curcio 1989), as well as more naturalistic treatment of scale (or features' size). The relatively simple model has proven very effective in predicting the visual fixations of an image viewer during task-free “free-viewing” (Itti and Koch 2000; 2001; Tatler et al. 2005), as high as between 84–98% according to one estimate (Harel et al. 2006). In practice, the computer model implementation of the model takes as an input an image (e.g., a cartoon) and outputs a saliency map specifying a saliency score for each location in the image (Fig. 1). An eye-tracker can record the locations within the image a participants' fixations fell on. The degree of overlap between model predictions and participants' fixations can then be computed—quantifying the degree to which a participant's gaze was guided by the image's salient features.

In the present investigation, professional improv comedians (improvs) and controls had their eye movements tracked while engaging in the task of captioning New Yorker cartoons. We have stipulated the following hypotheses:

H1: Improvs would gaze at salient cartoon features longer than controls.

H2: A longer duration of gaze at the salient features would result, on average, in funnier captions (as judged by independent raters).

Experiment 1: Eye-Tracking

Methods

Professional improv comedians and non-comedians improvised humorous and non-humorous captions to New Yorker cartoons while their eye movements were tracked.

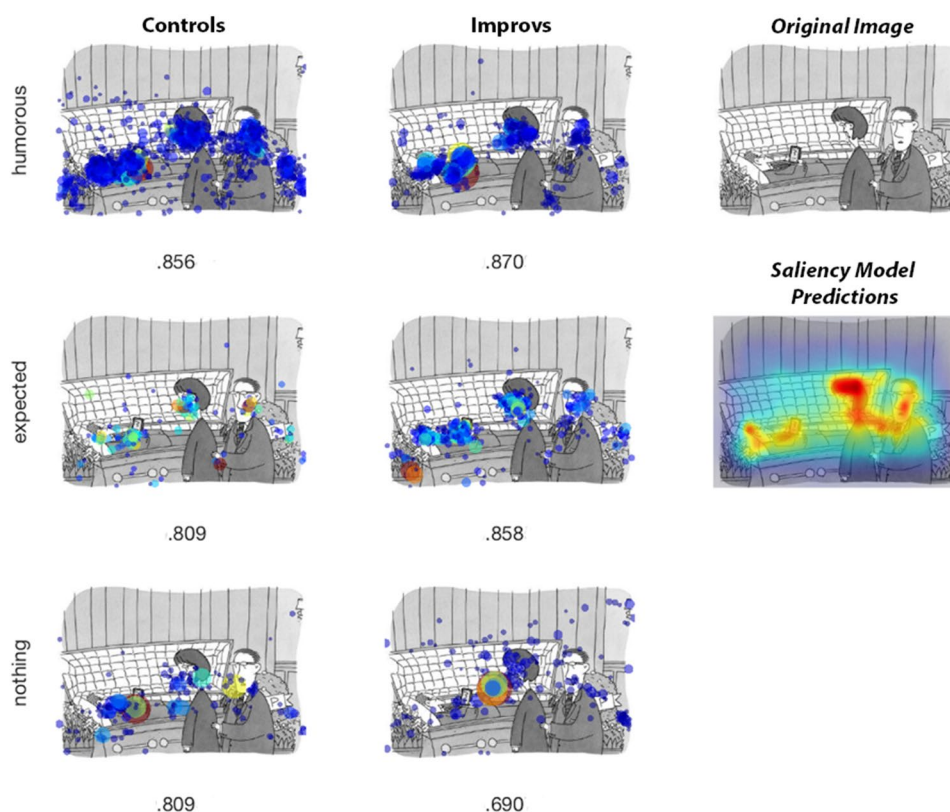
Participants

12 professional improv comedians and 12 controls participated in the experiment. The comedians (mean age 34.2; range: 26–52; 4 females) were contacted through their affiliations with various comedy clubs and professional improv groups. Four were members of the “Groundlings” improv group, while the remaining eight were affiliated with the “Second City” improv troupe. In our statistical analysis, all 12 comedians were combined into a single “improvs” group.

The majority of the controls (mean age 22.8; range: 18–48; 5 females) recruited were undergraduate students at XXXX College's Department of Psychology, who

² Anything that the joke is written about, e.g., a cartoon, improvised scene, news item.

Fig. 1 Fixations “heatmaps” for one of the cartoons used. Upper right: the image as the participants saw it (the cartoon is by Peter Vey © reprinted with permission). Middle right: saliency model predictions of eye-fixations based on low-level features. Left two columns: actual overall fixations of improv comedians and controls during the 3 experimental conditions (HUM, EXP, NOTH). Circles indicate visual fixations. Circle size indicates the total overall time spent fixating on a particular location. Circle color is meaningless, the purpose of varying the color is allowing a better understanding of the heatmaps when the fixation-circles partially overlap. Numbers under the heatmaps are ROC scores—the higher the score, the greater the fit between the model and the fixations



participated for course credit. A few additional controls were recruited from the College and were compensated monetarily. One control participant was a faculty member.

None of the participants were aware of the goals of the experiment while completing the tasks, and all had normal vision. XXXX College’s Institutional Review Board (IRB) approved all procedures of this experiment, and participants have electronically signed informed consent.

Stimuli and design

27 cartoons were selected from past issues of the New Yorker, with the captions and cartoonist signatures removed. 19 of the cartoons were black and white line drawings, and 8 used were multicolor.

Apparatus

The experiment was conducted on a 21.5” iMac running Windows 10 with Bootcamp. The experimental code was written in JavaScript and displayed with a Firefox web-browser on a locally hosted website. A Tobii 4C eye-tracker was mounted on the lower bezel of the iMac display and recorded the participants’ eye movements during the entire experiment.

Experimental procedure

Participants first electronically signed the informed consent, completed a demographic survey, and were given the instructions for the experiment. Participants were then guided through Tobii’s standard 9-point calibration process, using Tobii’s built-in software.

The experiment consisted of three tasks that dictated what type of caption should be generated by participants. The desired caption type was represented by a single word (Humorous, Expected, Nothing) that was presented for 2 s before each cartoon. During the Humorous (HUM) condition, participants were to generate a humorous caption that fit the cartoon image—specifically, a funny statement one of the cartoon characters would say in the situation. For Expected (EXP) trials, participants had to generate a non-humorous caption—a statement one may expect to hear in a non-funny situation. Finally, for the Nothing (NOTH) condition, participants were instructed to view the cartoon freely and not generate any caption.

Following the task prompt screen (2-s appearance of “Humorous”, “Expected”, or “Nothing”), a cartoon image was displayed. Participants were given a maximum of 30 s to come up with a caption for each image, and when time was up, the cartoon would be replaced by a page where they typed their captions and clicked “Next” to continue to the next cartoon (for HUM and EXP tasks only). Rather than

having to look at the cartoon for the full 30 s, if the participant felt they already had a caption in mind, they had the option of hitting the space bar to move on—that was in order to ensure that all eye movements recorded during the experiment were task related. Participants were instructed to observe the cartoon for the full 30 s during NOTH trials.

Each participant saw each of the 27 cartoons once, and the images were counterbalanced across the three conditions between participants. We also counterbalanced the order in which the cartoons were presented.

Data analysis

The heatmaps shown in Figs. 1 and 3 were generated using PyGaze (Dalmaijer et al. 2014). The centers of the circles are the estimated centers of the visual fixations. The sizes of the circles are proportional to the duration of the fixations. The color of the circles is meaningless and is meant to allow better understanding of the heatmap when the fixation-circles partially overlap. The heatmaps display all the fixations of a particular Group (Improvs/Controls) X Task (HUM/EXP/NOTH) combination during the first 10 s of visually scanning the cartoon—thus, the heatmaps display fixations of several participants within the same heatmap.

The main dependent variable of interest was the degree to which participants fixated on the visually salient features of the cartoons—salient in terms of low-level visual properties. To that end, we applied the Graph-Based Visual Saliency algorithm (Harel et al. 2007) to each cartoon image in MATLAB to generate saliency maps. These maps compute the most visually salient areas of images (e.g., Fig. 1, middle right image). Harel et al. (2007) found that, in the absence of top-down influences, these saliency maps predict much of the variance in human subjects' overt (i.e., revealed as eye movement) visual attention.

The degree of similarity between the saliency maps and the eye-tracking data was estimated using a receiver operating characteristics (ROC) score (using the same MATLAB toolbox of Harel et al.'s 2007). A ROC score of 1 signifies that the saliency map and the eye-tracking data overlap perfectly, while a ROC score of 0.5 indicates that the saliency map and the eye-tracking data are completely unrelated. ROC scores were computed separately for each image and each combination of Group (Improvs/Controls) X Task (HUM/EXP/NOTH) X two Timeframe windows: 1–5 s (defined as the full period between the instant the image was displayed and the moment 5 s have elapsed) and 6–10 s (from the instant 5 s have elapsed, i.e., the end of the first Timeframe, and until exactly 10 s of viewing the image concluded). Within each combination, all the relevant fixations were pooled together (e.g., pooling all fixations during the first 5 s of all improv comedians who were asked to generate

humorous captions to a particular cartoon) and compared to the saliency map of the cartoon.

Average ROC scores for each Group X Task X Timeframe combination were computed by averaging the ROC scores of that combination across the 27 images. Because the greatest source of variance—variance of no interest, were the differences between the images themselves, all statistical comparisons (e.g., improves vs. controls, or HUM vs. EXP) were made using repeated measures ANOVA or paired t-tests—paired with the same image.

Results

ROC scores

Professional improv comedians (“improvs”) and controls visually scanned cartoons while engaging in three cognitive tasks: thinking of a humorous caption (HUM), thinking of an expected, non-funny caption (EXP), and not attempting to think of any captions (NOTH). An overall average ROC score of 0.818 (range: 0.612–0.946) during the first 5 s indicated that low-level visual features' saliency predicted fixations much better than chance ($ROC=0.5$; $t(161)=70.0$, $p < <0.001$, Cohen's $d=5.5$), a typical range of scores for such images (Harel et al. 2007).

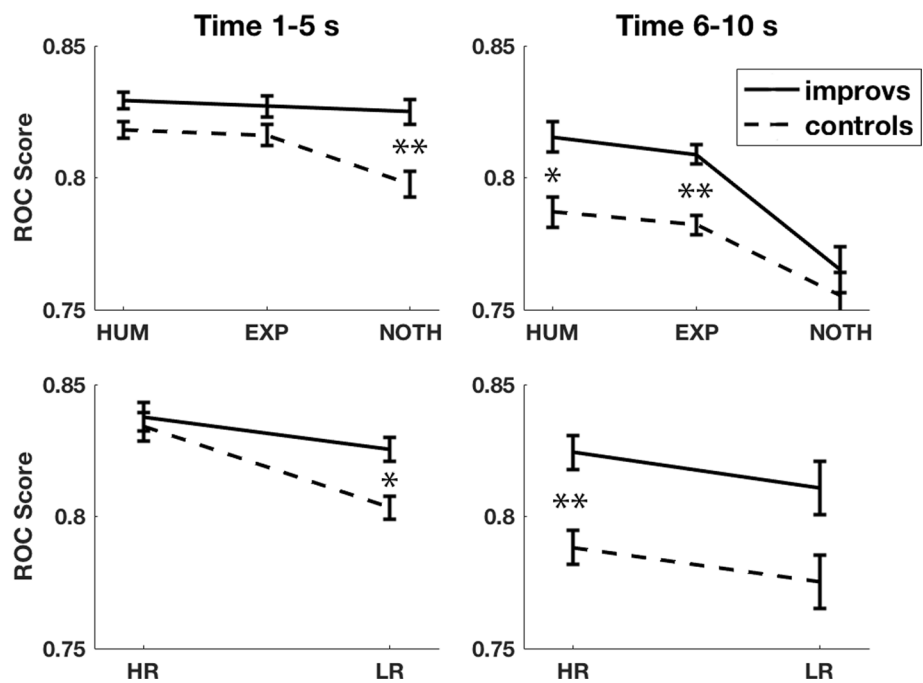
Task and expertise effects on ROC

A 3-way ANOVA (Task x Group x Timeframe) revealed significant main effects for each of the independent variables. A robust main effect for Group confirmed hypothesis H1, showing professional improv comedians' ROC scores ($M=0.812$, $SD=0.067$) were consistently higher than controls' ($M=0.793$, $SD=0.073$; $F(1,26)=14.7$, $p < 0.001$, $d=0.29$; during the first 5 s: $d=0.56$). As predicted, greater comedic experience or talent appeared to increase overt visual attention to salient cartoon features.

The main effect for task ($F(2,26)=9.5$, $p < 0.001$) revealed an increased rate of fixations on the salient image features with increased demands on creative output: HUM ROC rates ($M=0.812$, $SD=0.059$) as well as EXP's ($M=0.808$, $SD=0.060$) were greater than NOTH ROC rates ($M=0.786$, $SD=0.048$, $t(26)=3.32$ and 3.28 respectively, both $ps < 0.01$ (Bonferroni corrected for multiple comparisons), Cohen $ds=0.79$ and 0.54 , respectively). There was no significant difference between the HUM and EXP tasks ($t < 1$).

A main effect of Timeframe was also significant, with higher ROC scores in the first five seconds of visually scanning the cartoons ($M=0.825$, $SD=0.067$) than during seconds 6–10 ($M=0.799$, $SD=0.092$; $F(1,26)=7.3$, $p < 0.012$, $d=0.73$). The finding is in line with much of the saliency literature suggesting low-level visual features saliency effect

Fig. 2 Average ROC scores for improv comedians (solid lines) vs. controls (dashed lines) during the first 5 s of visually scanning the cartoons (left column) and seconds 6–10 (right column). The top plots illustrate the effect of task (generate a humorous caption [HUM], expected caption [EXP], or free-viewing [NOTH]) on ROC. The bottom plots illustrate the average ROC scores, within the HUM condition only, preceding the captions receiving the higher independent ratings (HR) vs those receiving the lower rating (LR)



on fixations reduces over time (e.g., Parkhurst et al. 2002) though, their influence does persist (Le Meur et al. 2007, 2006; Tatler et al. 2005).

An interaction between Task and Timeframe was significant ($F(2,52)=5.5$, $p<0.01$) and appears to be driven mostly by the more rapid decline of salient features' effect on fixations over time during the NOTH condition. A 3-way interaction Group x Task x Timeframe reached marginal significance ($F(2,52)=2.6$, $p<0.08$; see Fig. 2). No other interactions were found (all F s < 1).

A Mauchly test of Sphericity revealed a violation of the sphericity assumption of the repeated measures ANOVA for Task (Mauchly's $W=0.61$, $p<0.01$, Greenhouse–Geisser's $\epsilon=0.72$). However, applying a Greenhouse–Geisser correction did not qualitatively affect the results: the Task main effect remained similarly significant ($\epsilon=0.72$, adjusted $p<0.001$), and so was the Task x Timeframe interaction ($\epsilon=0.87$, adjusted $p<0.01$).

Experiment 2

Methods

To examine the relationship between visual fixation patterns and the quality (i.e., funniness) of the subsequently conceived captions, the cartoons along with participants' generated captions from Experiment 1 were presented to an independent group of participants to rate for funniness, cleverness, and surprise.

Participants

Seven undergraduate students from the Claremont Colleges were recruited to rate the HUM and EXP captions that were generated in Experiment 1. Participants were financially compensated.

Procedure

Using the same setup (21.5'' iMac, without the Tobii eye-tracker), each participant spent about 45 min rating a fraction (about half) of the total number of captions generated with a 7-point scale for funniness, cleverness, and surprise, with 1 representing the lowest possible score and 7 representing the highest (identical procedure to that of Amir and Biederman 2016). Each caption was rated by at least three different individuals.

Data analysis

Similar procedures as Experiment 1 were applied here to generate heatmaps and compute and compare ROC scores—however, the only eye-tracking data used in the present analysis were that obtained during HUM trials. HUM trials for each Experiment 1 participant were split into two conditions based on median independent raters' funniness ratings so that about half the trials were classified as High Rating (HR) and half as Low Rating (LR). Each participant in Experiment 1 generated 9 HUM captions (one-third of the 27 total trials). The nine captions were split into the two categories, HR and LR, by comparing each trial's caption funniness to

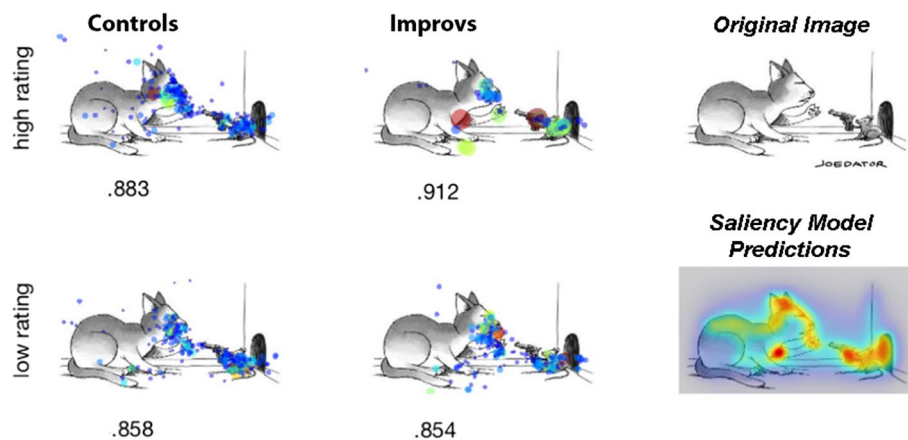


Fig. 3 Fixations “heatmaps” for one of the cartoons used. Upper right: the image as the participants saw it (minus the artist’s signature which was removed in the version displayed to the participants; the cartoon is by Joe Dator © reprinted with permission). Middle right: saliency model predictions of eye-fixations based on low-level visual features. Left two columns: actual overall fixations of improv comedians and controls during the trials that resulted in captions receiving the higher ratings (top) vs. lower ratings (bottom) all within the

humorous condition. Circles indicate visual fixations. Circle size indicates the total overall time spent fixating on a particular location. Circle color is meaningless, the purpose of varying the color is allowing a better understanding of the heatmaps when the fixation-circles partially overlap. Numbers under the heatmaps are ROC scores—the higher the score, the greater the fit between the model and the fixations

the participant’s median funniness rating; captions whose ratings were higher than the median funniness rating of the participant’s 9 HUM captions were categorized as HR, otherwise LR. Funniness Rating (HR vs. LR) was included in the ANOVA as an independent variable along with Group and Timeframe, with ROC scores as the dependent variable.

Results

First, we present the rating results themselves which serve as a manipulation check, e.g., confirming that professional improv comedians’ (improvs’) captions were rated as funnier than controls’ and that captions generated during the HUM condition were funnier than those of the EXP condition (see Rating Results). Subsequently, we explore the relationship between the pattern of visual fixation during caption generation and the caption’s funniness, confirming hypothesis H2, that increased duration of gaze at salient cartoon location results on average, in funnier captions (see ROC Results; for typical gaze pattern over one of the cartoons, see Fig. 3).

Rating results

A total of 438 captions (generated from the EXP and HUM tasks) were rated for their levels of funniness, surprisingness, and cleverness. 216 of the captions were written for the EXP task, while the remaining 222 were written for the HUM task. Each caption was rated by at least 3 different raters. Table 1 offers a summary of the overall rating data. Table 2 provides examples of captions written by improvs

and controls for the cartoon depicted in Fig. 1 along with the average ratings for each caption.

A 2-way ANOVA analysis of the Funniness ratings Group x Task reassuringly revealed a main effect of Task, namely that captions written during HUM trials were funnier than those written during EXP trials ($F(1,434) = 127.5$, $p < .001$, $d = 1.58$). An interaction of Group x Task ($F(1,434) = 3.9$, $p < .05$) was driven by the improvs’ funnier captions relative to controls’ during the HUM task ($t = 1.73$, $p < .05$, $d = 0.23$), while the two groups’ captions’ funniness were statistically indistinguishable during the EXP task ($t < 1$; see Table 1).

The measuring of surprise and cleverness was done for consistency with prior research (Amir and Biederman 2016) and was not hypothesis driven. The correlation between funniness and cleverness ($r = 0.80$, $p < .001$) and surprise ($r = 0.66$, $p < .001$) was quite high, and these measures are not discussed further here (but see supplementary information).

Table 1 Mean funniness, surprisingness, and cleverness ratings of EXP and HUM captions, for both subject groups. Standard deviation in parenthesis

Group	Condition	Funniness	Surprisingness	Cleverness
Controls	Expected	2.35 (.98)	2.58 (1.13)	2.36 (1.11)
	Humorous	3.32 (1.24)	3.81 (1.27)	3.42 (1.30)
Improvs	Expected	2.21 (.99)	2.01 (.99)	1.92 (.97)
	Humorous	3.59 (1.07)	3.68 (1.09)	3.57 (1.19)

Table 2 Example captions written by professional improv comedians (top) and controls (bottom) for the cartoon featured in Fig. 1, along with averaged independent ratings of their funniness, surprise, and cleverness

	Caption	Funniness	Surprise	Cleverness
Improvs	Melinda, honey, get over it. No one will read your novel	5	3.67	3.67
	Is it alright if I put my tiny hand on your back?	4.33	4	2.67
	Maybe now he'll actually finish the Song of Ice and Fire series	4.33	4.33	4.33
	He expired before his iphone	4	4	3.67
	He died doing what he loved—killing the book industry	2.67	3.33	2
Controls	Clearing one's browsing history	4.67	5.67	4.67
	No need to be sad, honey. This way, we can track his location, so we'll always know where he'll be	3.67	2.33	3.33
	Can dead bodies have boners?	3.33	5	3.33
	When we close it you'll lose the wifi, dad	3.33	4.33	4.33
	Shame about the crazy glue	2.33	2.67	2.33

ROC results

For the following analysis, the same eye-tracking data from Experiment 1 were used, but only data from the HUM task were included. In a 3-way ANOVA, the ROC scores were again the dependent variable, and, again, Group (Improvs/Controls) and Timeframe (1-5 s/6-10 s) were independent variables. However, in place of Task, the third independent variable was based on the captions' independent ratings: namely, whether the funniness rating was higher (HR) or lower (LR) than the median rating of the participant.

All three main effects were significant, with Group (ROC of improvs: $M = 0.82$, $SD = 0.08$, greater than controls': $M = 0.80$, $SD = 0.08$; $F(1,26) = 8.2$, $p < 0.01$, $d = 0.25$) and Timeframe (ROC for 1-5 s: $M = 0.83$, $SD = 0.07$, greater than 6-10 s: $M = 0.80$, $SD = 0.09$; $F(1,26) = 7.3$, $p < 0.02$, $d = 0.37$) revealing results similar to Experiment 1. In confirmation of hypothesis H2, an additional main effect of Funniness Rating revealed that greater visual fixations on the cartoon's visually salient features as measured by ROC resulted in funnier captions ($M = 0.82$, $SD = 0.07$) compared to captions conceived following reduced visual fixations on those features ($M = 0.80$, $SD = 0.09$; $F(1,26) = 5.1$, $p < 0.04$, $d = 0.25$).

Only one interaction reached marginal significance; Timeframe $\times$ Group ($F(1,26) = 3.3$, $p < 0.08$) appears to be driven by the controls losing attention to the salient features faster than the improvs (Fig. 2, bottom panels).

Discussion

Professional improv comedians and controls engaged in cartoon captioning task while their eye movements were being tracked. We have hypothesized that increased reliance on bottom-up processing, that is, allowing the cartoons to

generate spontaneous associations as opposed to a more top-down controlled creative process, would be the hallmark of comedic expertise. The hypothesis was based on Amir and Biederman's (2016) brain imaging experiment in which comedians showed greater activation in the semantic regions of the temporal cortex and lesser activity in the prefrontal cortex, while engaging in humor creativity, relative to controls. As an index of bottom-up processing, we used a measure of the degree to which participants fixated at the cartoons' salient features—a measure obtained by Harel et al.'s (2007) saliency model. Specifically, we hypothesized that the comedians would gaze at salient cartoon features more than controls, and that, within participant, an increased rate of fixation at salient cartoon features would lead to funnier captions. Both hypotheses were robustly confirmed.

We summarize the main findings as follows:

1. Professional improv comedians fixated at salient cartoon locations more than controls.
2. The task of generating a caption resulted in increased fixations at salient cartoon locations relative to free-viewing.
3. An increased visual attention to the cartoons' salient features resulted, on average, in a funnier caption.

Saliency as an index for bottom-up processing

In free-viewing tasks, the saliency model (Itti et al. 1998; Harel et al. 2006) can predict much of the variance in observers' visual fixations based on simple, low-level visual features alone (e.g., as simple as local contrasts in luminance, color, and edge orientation). We found that increased fixation on such low-level salient features during the humor captioning task of cartoons was greater in the professional improv comedians group compared with controls. Increased attention to the salient cartoon regions positively correlated

with the subsequently generated caption's funniness. These results may appear counterintuitive at first glance: why should success with the high-level cognitive process of comedic creativity be linked to increased attention to the low-level prompt features? We believe this apparent incongruity may be resolved by interpreting increased fixation on the salient features, in the context of this study's task, as an index of increased bottom-up processing—which in turn may be more effective in producing cartoon *relevant* captions.

Most, if not all visual tasks, involve top-down as well as bottom-up image processing (Wolfe et al. 2007; Melloni et al. 2012). A task may shift attentional bias, resulting in increased tendency to fixate upon particular locations or at a particular type of feature—thus, departing from the predictions of the saliency models (Navon and Kasten 2011; Harding and Robertson 2009). Thus, while free-viewing overt visual attention is largely directed by an image driven, bottom-up process (as suggested by the finding that 84–98% of fixations during free-viewing procedures can be predicted by simple features of the image itself; Harel et al. 2006), the introduction of a task may result in top-down control over the visual scanning, competing with the salient features of the image itself in directing overt visual attention (i.e., eye movements). Keeping the image constant, a greater degree of accuracy in the prediction of visual fixations could thus serve as an index to the degree to which bottom-up processing is employed. With that interpretation in mind, we may interpret our findings to suggest:

1. professional improv comedians engage in more bottom-up processing than controls,
2. the task of generating a caption elicits greater bottom-up processing than free-viewing, and
3. a greater degree of bottom-up processing results, on average, in a funnier caption.

It is important to note that while the extent of bottom-up processing may be indexed by the portion of the variance in visual saccades explained by predictions of the saliency model (Harel et al. 2007), the unexplained variance may be the result of at least two causes: (a) a competing top-down process directing attention away from salient image features and toward a more task relevant and less salient features. (b) An overall decreased attention or cognitive processing dedicated to the image as well as the task (if there is a task). With the measures employed in an eye-tracking experiment, it is not possible to conclusively differentiate the two explanations, but only (a) is consistent with evidence from previous work, which showed, for example, that overall activation of the brain regions involved in comedy writing did not vary between comedians and controls during a similar task, but rather the proportion of bottom-up vs. top-down activation

did (such that comedians displayed a greater proportion of bottom-up activation, and decreased top-down interference increased captions funniness; Amir and Biederman 2016).

Captioning task effects

Unlike well-defined tasks which may result in shifting gaze from the most salient image features to task relevant features (e.g., as in a task to scan an image in search for cups; Harding and Robertson 2009), the creative task of generating a caption appears to increase the degree to which participants' gaze was focused on the salient features, relative to free-viewing the image. Since the task of free-viewing of the cartoon is unlikely to engage greater top-down processing, as there is no goal to direct visual exploration toward, the most likely explanation for the finding is increased attention to and processing of the image overall in the humorous captioning condition. Indeed, in a neuroimaging study of the same task, Amir and Biederman (2016) found that, during humorous captioning trials, greater activation appeared in both bottom-up (e.g., TMP) and top-down (e.g., mPFC) processing regions compared to both the mundane captioning condition and task-free cartoon viewing. In other words, the creative task of humor captioning requires increased cognitive processing overall—the bottom-up component of which is revealed as increased rate of fixation at salient features.

Effect of comedic expertise

Professional improv comedians have consistently across all conditions gazed at the salient image features more than controls. This result, at first glance, may appear surprising, since the predictions of the saliency model (Harel et al. 2007) are based entirely on local simple features, and comedic expertise of improv comedians is not, presumably, based on superior visual processing—rather, more high-level semantic skills (Greengross et al. 2012). Indeed, when a *visual* task is involved, experts often gaze away from the salient features in favor of the task relevant features (Gegenfurtner et al. 2011; e.g., experts vs. novice physicians spend more time gazing at polyps while viewing colonoscopy images; Bernal et al. 2014), however, in some domains, experts do rely on the same features as novices (e.g., when differentiating among birds of the same subspecies; Amir et al. 2011). As discussed above, we suggest the increased overt attention to salient features indexes an increase in bottom-up processing. When performing the same captioning task in the MRI, professional comedians' semantic regions in the temporal lobes showed increased activation relative to amateurs, and amateurs' activity there was greater than controls'—the reverse was true for the mPFC (Amir and Biederman 2016; Brawer and Amir 2021). The authors interpreted the results to suggest an increased bottom-up processing of the cartoons

and decreased top-down (or cognitive control) processing with increased comedic expertise. Our eye-tracking results are consistent with their hypothesis.

Funniness ratings

All of the captions produced during the eye-tracking study were subsequently rated by independent raters (see experiment 2). Reassuringly, captions composed during the Humorous condition were rated as much funnier than those composed during the Expected condition, and comedians' captions were rated as funnier than controls'. We found that participants' eye movements predicted, on average, the funniness of the caption they ended up generating. Whether comedians or controls, the funnier captions were produced following an increased gaze rate upon the salient features. As with expertise, we interpret the result to suggest a greater degree of bottom-up processing results, on average, in a greater quality of the creative product (i.e., greater funniness). Indeed, a common guidance in improv comedy classes is to “get out of one's head” (decrease top-down processing), and to listen and let the scene inspire the reaction (bottom-up processing)—such a cognitive strategy should result in a response more *relevant* to the improvised scene/cartoon. Indeed, the funniest captions in the New Yorker cartoon captioning were those involving *relevant* associations that were not the most obvious, yet not too obscure (Shahaf et al. 2015). To find such associations, one may be best served by letting the cartoon drive their flow of association maximally and in a bottom-up fashion. In support of this interpretation, Amir and Biederman (2016) found that increase in temporal association regions activity was associated with funnier captions—activity they suggest reflected a bottom-up search for remote semantic associations which formed the basis for a creative humorous caption (Houston and Mednick 1963).

Gaze over time

Participants' gaze's overlap with salient image regions reduced over time regardless of condition or comedic expertise. Indeed, the effect of saliency on gaze while persistent is often reduced as a function of the duration of time of viewing an image (Parkhurst et al. 2002; Le Meur et al. 2007, 2006; Tatler et al. 2005). In the present experiment, the rate of reduction in gaze predictability by the saliency map is inversely related to task demand—with slowest reduction during the HUM condition, fastest during NOTH (i.e., free-viewing). Some investigators of creativity suggested a two-step process of creativity: ideation followed by evaluation (e.g., Basadur 1995). Ideation is more of a bottom-up process, while evaluation is more of a top-down as reflected in the brain activity associated with the two stages (Shah et al. 2013). The reduction of attention to salient features

over time may reflect a switch from ideation (bottom-up) to evaluation (top-down) processing—in addition to other potential factors normally causing a reduction of saliency impact on gaze over time (such as overall reduction in attention to the image, as well as “inhibition of return”; see Wang and Klein 2010).

Comedy writing vs. general creativity

While the current investigation focused on *comedy* writing and demonstrated that increased bottom-up processing improves the quality of the creative product (i.e., caption funniness) and is associated with expertise in the creative domain (i.e., being a professional comedian), there are reasons to speculate that the conclusion may be generalized to other creative domains. Unlike with practical problem solving in which the goal is typically well defined (e.g., find one's son in the crowd), creative tasks are characterized by vaguer goals (e.g., write a caption that's funny, make a drawing that's beautiful). A vague goal may not be as effective at competing with the prompt's salient features since the goal is high-level, complex, and abstract and thus harder to translate to a preference for low-level features to be attended to, other than the salient ones—thus, low-level saliency triumphs and drives the creative process in a bottom-up fashion. Similar to Amir and Biederman's (2016) finding that professional comedians displayed decreased prefrontal activation during comedy writing, Pinho et al. (2014) found a decrease in prefrontal activation to be associated with increased musical expertise, indicating that both creative endeavors require reduced top-down processing with increased experience. We proposed the reason bottom-up processing is effective in comedy writing is because it can result in associations that are both relevant and remote/unexpected given a prompt. Mednick's (1962) popular thesis proposed that the basis for creativity in general is the meaningful (read: *relevant*) linking of *remote* associations. Note that our results only suggest that experts engage in a higher *degree* of bottom-up processing. We believe both bottom-up and top-down processing play a role in creative endeavors (e.g., as evidenced by the prefrontal activation reported by most imaging studies of creativity; Dietrich and Kanso 2010).

Beaty et al. (2014) proposes a distinction between the *associative* and *controlled attention* theories of creative thought. The associative theory sees creativity as a largely *bottom-up* process in which associations spread spontaneously until a creative idea or insightful solution is reached, and that more creative individuals have “flatter” associative hierarchies supporting more remote associations (Mednick 1962; Bowden et al. 2005). Associative theories are consistent with the results of the current investigation and are further supported by decreased prefrontal activation in professional comedians and musicians relative to controls

during improvisation (Amir and Biederman 2016; Pinho et al. 2014). Controlled attention theories posit that the key to creativity is the *top-down* interference in the spontaneous flow of associations, by inhibiting³ the more frequent obvious associations and thus allowing the more remote associations consideration in the creative process (Beaty et al. 2014; Frith et al. 2021). Support for the controlled attention theories includes the activation found in prefrontal cortex in nearly all creative endeavors (Dietrich and Kanso 2010), the increased functional connectivity of the executive control and saliency neural networks in creative individuals (Beaty et al. 2018), and the positive correlation between attentional control and creative abilities (Frith et al. 2021). In all likelihood, both top-down and bottom-up mechanisms play a role and the question is one of degree. For example, it has been proposed that the creative process could be subdivided into a largely bottom-up *idea generation* stage followed by a top-down *selection* stage (Shah et al. 2013). Such two-stage characterization is consistent with our finding of peak bottom-up processing at the beginning of the trial—as indexed by greater ROC scores—followed by a gradual decline of the measure.

Future research would determine if indeed an increased proportion of bottom-up processing is associated with expertise in other creative domains, and whether it typically results in a higher quality creative product.

Summary and conclusion

A map of the salient low-level image features does better at predicting gaze during the improvisation of humorous captions compared to free-viewing. The extent of gazing at salient features correlate positively with comedic expertise and the funniness of the captions. We propose the predictability of gaze by the saliency map reflects the extent of bottom-up processing of the cartoons. Such increase in stimulus-driven creativity is a more effective strategy to generate funny ideas perhaps since it is the most effective way to generate remote and unexpected associations/interpretations/perspectives that are nevertheless stimulus relevant. To be funniest, one must address the elephant in the room—from an unexpected perspective.

³ Note that a process of the inhibition of previously activated associations may occur in the absence of top-down intervention (Martindale 2007). This extends to overt attention: in eye-tracking studies in which participants engage in a task free, free viewing of images, participants often do not redirect their gaze into image locations they recently scanned—a phenomenon known as inhibition of return (Wang & Klein, 2010).

[Box 1:] advice to aspiring comedians

Rather than suggesting ground breaking advice, our results offer independent evidence for age-old comedy coaches' wisdom:

- Pay attention to the salient features of the prompt (situation/cartoon/remark) so that your jokes would be more relevant (or in comedy coaches' jargon: "address the elephant in the room").
- Allow the situation to inspire comedic creativity in a bottom-up fashion ("listen and react").
- Reduce top-down impositions on the creative process, i.e., excessive planning or searching for humor outside of the present situation or prompt ("get out of your head").

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1007/s10339-022-01079-0>.

Funding The research was supported by an NSF grant IIS-1948517 to AP and a Pomona College Faculty Grant to OA.

Data availability The datasets generated during and/or analyzed during the current study are available from the corresponding author on reasonable request.

Declarations

Conflict of interest On behalf of all authors, the corresponding author states that there is no conflict of interest.

References

- Amir O (2016) The frog test: a tool for measuring humor theories' validity and humor preferences. *Front Hum Neurosci* 10:40
- Amir O, Biederman I (2016) The neural correlates of humor creativity. *Front Hum Neurosci* 10:597
- Amir O, Xu X, Biederman I (2011) The spontaneous appeal by naïve subjects to nonaccidental properties when distinguishing among highly similar members of subspecies of birds closely resembles descriptions produced by experts. In: Vision sciences society annual meeting
- Amir O, Biederman I, Wang Z, Xu X (2013) Ha Ha! Versus Aha! A direct comparison of humor to nonhumorous insight for determining the neural correlates of mirth. *Cereb Cortex* 25(5):1405–1413
- Aziz-Zadeh L, Liew SL, Dandekar F (2013) Exploring the neural correlates of visual creativity. *Soc Cogn Affect Neurosci* 8(4):475–480
- Basadur M (1995) Optimal ideation-evaluation ratios. *Creat Res J* 8(1):63–75
- Beaty RE, Silvia PJ, Nusbaum EC, Jauk E, Benedek M (2014) The roles of associative and executive processes in creative cognition. *Mem Cognit* 42(7):1186–1197
- Beaty RE, Kenett YN, Christensen AP, Rosenberg MD, Benedek M, Chen Q, Silvia PJ (2018) Robust prediction of individual creative ability from brain functional connectivity. *Proc Natl Acad Sci* 115(5):1087–1092

- Bernal J, Sánchez FJ, Vilariño F, Arnold M, Ghosh A, Lacey G (2014) Experts vs. novices: applying eye-tracking methodologies in colonoscopy video screening for polyp search. In: Proceedings of the symposium on eye tracking research and applications, pp 223–226
- Bowden EM, Jung-Beeman M, Fleck J, Kounios J (2005) New approaches to demystifying insight. *Trends Cogn Sci* 9:322–328
- Brawer J, Amir O (2021) Mapping the ‘funny bone’: neuroanatomical correlates of humor creativity in professional comedians. *Soc Cogn Affect Neurosci*
- Chen Q, Beaty RE, Cui Z, Sun J, He H, Zhuang K, Qiu J (2019) Brain hemispheric involvement in visuospatial and verbal divergent thinking. *Neuroimage* 202:116065
- Chen Q, Beaty RE, Qiu J (2020) Mapping the artistic brain: common and distinct neural activations associated with musical, drawing, and literary creativity. *Hum Brain Map*
- Dalmaijer ES, Mathôt S, Van der Stigchel S (2014) PyGaze: an open-source, cross-platform toolbox for minimal-effort programming of eye tracking experiments. *Behav Res Methods* 46:913–921
- Dietrich A, Kanso R (2010) A review of EEG, ERP, and neuroimaging studies of creativity and insight. *Psychol Bull* 136:822–848
- Erhard K, Kessler F, Neumann N, Ortheil HJ, Lotze M (2014) Professional training in creative writing is associated with enhanced fronto-striatal activity in a literary text continuation task. *Neuroimage* 100:15–23
- Frith E, Kane MJ, Welhaf MS, Christensen AP, Silvia PJ, Beaty RE (2021) Keeping creativity under control: contributions of attention control and fluid intelligence to divergent thinking. *Creat Res J* 1–20
- Gegenfurtner A, Lehtinen E, Säljö R (2011) Expertise differences in the comprehension of visualizations: a meta-analysis of eye-tracking research in professional domains. *Educ Psychol Rev* 23(4):523–552
- Greengross G, Martin RA, Miller G (2012) Personality traits, intelligence, humor styles, and humor production ability of professional stand-up comedians compared to college students. *Psychol Aesthet Creat Arts* 6(1):74
- Harding P, Robertson NM (2009) A comparison of feature detectors with passive and task-based visual saliency. In: Scandinavian conference on image analysis, pp 716–725. Springer, Berlin
- Harel J, Koch C, Perona P (2007) Graph-based visual saliency. *Adv Neural Inf Process Syst* 19:545–552
- Harel J, Koch C, Perona P (2006) Graph-based visual saliency. *Advances in neural information processing systems*, 19
- Hirsch J, Curcio CA (1989) The spatial resolution capacity of human foveal retina. *Vis Res* 29(9):1095–1101
- Houston JP, Mednick SA (1963) Creativity and the need for novelty. *Psychol Sci Public Interest* 66(2):137
- Howard-Jones PA, Blakemore SJ, Samuel EA, Summers IR, Claxton G (2005) Semantic divergence and creative story generation: an fMRI investigation. *Cogn Brain Res* 25:240–250
- Hummel JE, Biederman I (1992) Dynamic binding in a neural network for shape recognition. *Psychol Rev* 99(3):480
- Hurley MM, Dennett DC, Adams RB Jr., Adams RB (2011) *Inside jokes: using humor to reverse-engineer the mind*. MIT Press
- Itti L, Koch C (2000) A saliency-based search mechanism for overt and covert shifts of visual attention. *Vis Res* 40(10–12):1489–1506
- Itti L (2000) Models of bottom-up and top-down visual attention. California Institute of Technology
- Itti L, Koch C (2001) Computational modelling of visual attention. *Nat Rev Neurosci* 2(3):194–203
- Itti L, Koch C, Niebur E (1998) A model of saliency-based visual attention for rapid scene analysis. *IEEE Trans Pattern Anal Mach Intell* 20(11):1254–1259
- Itti L (2000) Models of bottom-up and top-down visual attention. California Institute of Technology
- Kinchla RA, Wolfe JM (1979) The order of visual processing: “Top-down”, “bottom-up”, or “middle-out.” *Percept Psychophys* 25(3):225–231
- Le Meur O, Le Callet P, Barba D, Thoreau D (2006) A coherent computational approach to model the bottom-up visual attention. *IEEE Trans Pattern Anal Mach Intell* 28:802–817
- Le Meur O, Le Callet P, Barba D (2007) Predicting visual fixations on video based on low-level visual features. *Vis Res* 47(19):2483–2498
- Limb CJ, Braun AR (2008) Neural substrates of spontaneous musical performance: an fMRI study of jazz improvisation. *PLoS ONE* 3:e1679. <https://doi.org/10.1371/journal.pone.0001679>
- McGraw AP, Warren C (2010) Benign violations: Making immoral behavior funny. *Psychol Sci* 21(8):1141–1149
- Martin RA, Ford T (2018) The psychology of humor: an integrative approach. Academic press
- Martindale C (2007) Creativity, primordial cognition, and personality. *Personality Individ Differ* 43(7):1777–1785. <https://doi.org/10.1016/j.paid.2007.05.014>
- Mednick S (1962) The associative basis of the creative process. *Psychol Rev* 69(3):220. <https://doi.org/10.1037/h0048850>
- Melloni L, van Leeuwen S, Alink A, Müller NG (2012) Interaction between bottom-up saliency and top-down control: how saliency maps are created in the human brain. *Cereb Cortex* 22(12):2943–2952
- Navon D, Kasten R (2011) A demonstration of direct access to colored stimuli following cueing by color. *Acta Physiol (oxf)* 138(1):30–38
- Parkhurst D, Law K, Niebur E (2002) Modeling the role of salience in the allocation of overt visual attention. *Vis Res* 42(1):107–123
- Pinho AL, de Manzano O, Fransson P, Eriksson H, Ullen F (2014) Connecting to create: expertise in musical improvisation is associated with increased functional connectivity between premotor and prefrontal areas. *J Neurosci* 34(18):6156–6163
- Raskin V (2012) *Semantic mechanisms of humor*, Vol 24. Springer
- Rominger C, Papousek I, Perchtold CM, Benedek M, Weiss EM, Weber B, Fink A (2020) Functional coupling of brain networks during creative idea generation and elaboration in the figural domain. *NeuroImage*, 207
- Ruch W (2001) The perception of humor. In: *Emotions, qualia, and consciousness*, pp 410–425
- Ruch W, Attardo S, Raskin V (1993) Toward an empirical verification of the general theory of verbal humor. *Humor Int J Humor Res* 6(2):123–136
- Schlegel A, Alexander P, Fogelson SV, Li X, Lu Z, Kohler PJ et al (2015) The artist emerges: visual art learning alters neural structure and function. *Neuroimage* 105:440–451
- Shah C, Erhard K, Ortheil HJ, Kaza E, Kessler C, Lotze M (2013) Neural correlates of creative writing: an fMRI study. *Hum Brain Mapp* 34(5):1088–1101
- Shahaf D, Horvitz E, Mankoff R (2015) Inside jokes: identifying humorous cartoon captions. In: *Proceedings of the 21th ACM SIGKDD international conference on knowledge discovery and data mining*, pp 1065–1074
- Suls JM (1972) A two-stage model for the appreciation of jokes and cartoons: an information-processing analysis. *Psychol Humor Theoret Perspect Empir Issues* 1:81–100
- Tatler BW, Baddeley RJ, Gilchrist ID (2005) Visual correlates of fixation selection: effects of scale and time. *Vis Res* 45(5):643–659
- Theeuwes J (2010) Top-down and bottom-up control of visual selection. *Acta Physiol (oxf)* 135(2):77–99
- Villareal MF, Cerquetti D, Caruso S, Schwarcz Lopez Aranguren V, Gerschovich ER, Frega AL, Leiguarda RC (2013) Neural correlates of musical creativity: differences between high and low creative subjects. *PLoS ONE* 8(9):75427

- Wolfe JM, Horowitz TS, Van Wert MJ, Kenner NM, Place SS, Kibbi N (2007) Low target prevalence is a stubborn source of errors in visual search tasks. *J Experim Psychol :General* 136(4):623
- Wang Z, Klein RM (2010) Searching for inhibition of return in visual search: a review. *Vis Res* 50(2):220–228

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.