# A Neuroscientist Walks Into a Bar: Our Insights on Humor or the Lack Thereof

By Hardik Sangwan

Humor, maybe more so than anything else, highlights how little we comprehend our own minds. Said to be the most complex problem/trait to replicate and understand in both AI and cognitive science, humor is our brain at it’s finest; what’s funny can be so simple yet so complex. It is so inherently related to one’s experiences and understanding that it seems only a perfect duplication of someone’s brain could help us decode what they would find funny. Not much exists in terms of theories of humor, especially ones that are generally agreed upon by the major scientific community or ones that even cover all the different cases of humor. [1] Two major theories from the past are the Superiority Theory, which simply means that humor indicates a sense of superiority over others/our own previous selves, and the Incongruity Theory, which states that humor is a result of a violation of our expectations. And while the Incongruity Theory is an improvement over the Superiority Theory, it too fails to explain a number of cases such as when something is funny to one person but makes another cringe, or when we feel an emotion other that humor as a result of the incongruity. Other noteworthy theories not analyzed in this paper are the Benign Violation Theory [2] which states that humor occurs when we experience a benign violation of our morals and the General Theory of Verbal Humor [3] which describes humor though parameters called Knowledge Resources so any statement can be provided a score based on these ‘KRs’. Similar ideas of surprise/violation and resolution, are evident in all these theories and the others to be discussed further.

This paper will explore relatively recent theories that have been proposed to provide us with a better understanding of humor while also resolving it in such a way that it complies with modern computing technologies and cognitive science. The papers involve both experimental and descriptive studies by the authors to get a better grasp of this problem.

The latest major theoretical development in the field seems to have come in the form of a quantum theory of humor [4] that bases the cognitive properties of humor in a similar mathematical context to that of the realm of quantum physics; so possible interpretations are treated as linear super positions of basis states in a Hilman space, meaning that the mathematical framework is used to create probability models that essentially can collapse into funny or not funny. The theory is influenced by previous studies of psychological phenomena using quantum models, ranging from word processing to vision to memory. In classical probability models events come from a common sample space, whereas quantum models of probability define states and variables with reference to a context. This method of collapsing into a definite state from context based ambiguity is similar to the ambiguity often involved with humor and therefore the quantum approach naturally lends itself to the study of humor. Consider the classic example given in the paper,

“Time flies like an arrow; fruit flies like a banana”

The words ‘flies’ and ‘like’ here have different conflicting meanings that have to be resolved to understand the statement and that resolution lies at the core of humor. So processing the multiple meanings provided by the words can be considered the ‘quantum superposition’ and then the superposition collapses when the ‘location’ of the words is measured or the meaning is understood. To prove that humor does not follow the classical model of probability, the authors conducted an experiment on 85 participants with 35 jokes and variants to test their theory. Each joke has multiple variants that obtain a funniness rating through the participants and then an average funniness value for the joke can be calculated. According to the classical probability theory, this value should be the same as the sum of funniness of all variants. If that is not the case, then it could be argued that the classical model does not apply and quantum models might suit this problem better; assuming that the experiment is correct. And while the experiment showed that the sums and average do not add up, to a lot of researchers, this is most likely a result of an inaccurate/incomplete experiment. Even so, the approach does seem promising in terms of providing a path to use our computational abilities to understand jokes.

Another interesting approach, by Boyang Li from Georgia Tech, presents a Dynamic and Dual Process Theory of Humor [5]. Li tries to provide a framework for how humor works based on dual process cognition and emotional appraisal, stating that it can be described as a four step dynamic process consisting of surprise, reflection, dismissal and compensation. Humor is treated, not as an independent subsystem, but in conjunction with other emotional appraisals. Dual process simply refers to the idea that cognitive processes are either automatic and fast or deliberate and slow. The first step in the dynamic process is surprise, wherein the situation is confusing enough to engage the deliberate process of cognition and not rely on intuitive automatic processing. ‘A quick realization that the surprise is not worthy of mental effort’ leads to a positive emotional response. Therefore the dismissal mechanism, the third step of the dynamic process, becomes crucial in identifying whether the situation will be considered funny or threatening or maybe a learning opportunity. For example, consider the puzzle below

*What walks on four legs in the morning, two in the afternoon, three in the evening and no legs at night?*

The surprise comes from not knowing the answer. Once the answer is known, the surprise is resolved. But the riddle is not funny because the realization is of learning something new and therefore it’s purpose is served and no dismissal step happens. Li argues that the dismissal is crucial to humor since surprise is at least a slightly negative emotion, and a quick dismissal of the surprise after reflection is what leads to a quick change from negative to positive emotion which in turn leads to instantaneous humor.

A more detailed analysis of incongruity-resolution/surprise-dismissal was conducted by Sascha Topolinski [6] to obtain a relationship between the fluency of processing and funniness. Processing fluency is simply the easiness with which an insight occurs. Topolinski applies this approach to humor processing to argue that the ‘easier’ a resolution is, the more positive effect it has and therefore the funnier it becomes. In his paper, Topolinski conducts multiple experiments which manipulate fluency using change of type font to change the encoding and change in the amount of repetition. In one of the experiments, the researcher pre exposes punchline words at different times before a joke to see the effects on funniness, surprisingly finding that preexposing a punchline up to a minute before the actual joke can increase funniness but exposure right before the joke decreased funniness. In case of the type font experiment, it is made clear that the easier type font leads to a higher funniness rating from the subjects of the trial. Topolinski further argues that non verbal humor could also be studied similarly using non-semantic fluency variations such as perceptual fluency or motor fluency.

These theories and the understanding of humor in general can relate back to almost all the topics we have studied in our book so far, but one particular area that stands out is Linguistics [7]. Phonetic and semantic patterns become important for understanding jokes along with the differences due to the structure of the language being used i.e., English or French or Spanish. And just how we have theorized a universal grammar for language, a similar universal basis is likely for humor that applies to any given language. And as we make strides in natural language processing especially as it relates to other emotions like sadness or anger, the progress directly relates to potential progress in our understanding of humor as well. It is the ‘pragmatic analysis’ stage of NLP that poses the biggest issue when trying to understand humor. Models such as the Wernicke-Geschwind could be used along with a brain imaging method to test which particular areas of the brain light up when we are presented with a funny joke vs a neutral statement. According to the model, a joke heard or ‘seen’ will be processed in the either in primary auditory cortex or the visual cortex, but then is ultimately passed to Wernicke’ area where the key processing happens and essentially an understanding is born; therefore we would have to look at specifically this region to find out differences between humorous and non humorous scenarios. Another area of study that follows directly from this is to see if subjects with brain damage in this area can still understand humor or not; and specifically, are there particular types of humor they stop understanding because of the damage. While we know through modern imaging techniques that there are other regions of the brain that contribute to language function as well, the general idea of trying to find specific regions which handle language and then analyzing them under the humor context still applies.

While the study of humor has been gaining more and more traction over the recent past, much remains to be done especially in our ability to create a ‘funny’ AI agent or even one that can understand what is funny or not. There have been multiple papers[8][9][10] that use specific constrained situations such as focusing on only puns or only memes to identify and recreate that type of humor using some form of machine learning/classification. And beyond the theories discussed here, humor can be a lot more than just the meaning of words; we can find sounds funny, or something visual or some combination of things together can trigger a laugh. Therefore the problem becomes even more complex when we include all the different contexts where we find humor. We seem to only be at the stage of ‘sidewalk neuroscience’ in terms of it, meaning mostly simple observational procedures without a concrete theory. It is likely that humor will be the last frontier in our scientific endeavors on our own minds.. At the same time, research that uses the latest technologies like fMRIs and PET scans along with the capabilities of Big Data analysis and Deep learning seems like a promising approach into the field. The search is on for a Universal Theory of Humor.

**References**

1. Morreall, John, "Philosophy of Humor", *The Stanford Encyclopedia of Philosophy*(Winter 2016 Edition), Edward N. Zalta (ed.), URL = <https://plato.stanford.edu/archives/win2016/entries/humor/>.

2. McGraw, A. P., & Warren, C. (2010). Benign Violations: Making Immoral Behavior Funny. *Psychological Science*, *21*(8), 1141–1149. <https://doi.org/10.1177/0956797610376073>

3. Humor - International Journal of Humor Research, Volume 24, Issue 2, Pages 203–222, ISSN (Online) 1613-3722, ISSN (Print) 0933-1719, DOI: <https://doi.org/10.1515/HUMR.2011.013>

4. Gabora, L., & Kitto, K. (2017). Toward a Quantum Theory of Humor. *Frontiers in Physics,* *4*. doi:10.3389/fphy.2016.00053

5. Li, Albert (2015). A Dynamic and Dual Process Theory of Humor. Proceedings of the Third Annual Conference on Advances in Cognitive Systems

6. Sascha Topolinski (2014) A processing fluency-account of funniness: Running gags and spoiling punchlines, Cognition and Emotion, 28:5, 811-820, DOI: [10.1080/02699931.2013.863180](https://doi.org/10.1080/02699931.2013.863180)

7. Friedenberg, J., & Silverman, G. (2016). The Linguistic Approach: Language and Cognitive Science. In *Cognitive Science: An Introduction to the Study of Mind* (2nd ed., pp. 275-310). Thousand Oaks: SAGE Publications.

8. Reyes , A., Rosso , P. & Buscaldi , D. (2011). Humor in the Blogosphere: First Clues for a Verbal Humor Taxonomy. *Journal of Intelligent Systems*, 18(4), pp. 311-332. Retrieved 5 Apr. 2019, from doi:10.1515/JISYS.2009.18.4.311

9. Mihalcea, R. and Strapparava, C. (2006), LEARNING TO LAUGH (AUTOMATICALLY): COMPUTATIONAL MODELS FOR HUMOR RECOGNITION. Computational Intelligence, 22: 126-142. doi:[10.1111/j.1467-8640.2006.00278.x](https://doi.org/10.1111/j.1467-8640.2006.00278.x)

10. Raz, Yishay (2012), Automatic humor classification on Twitter. Association for Computational Linguistics, pp. 66-70.

**PRINT AND ATTACH TO YOUR PAPER**

**AFTER REFERENCE PAGE**

**Student # 903075634 Word Count 2000**

**Paper Rubric**

|  |  |  |
| --- | --- | --- |
|  | **Points POSSIBLE** | **Points EARNED** |
| **Introduction** | **10** |  |
| **Argument/Implications** | **20** |  |
| **Examples** | **15** |  |
| **Conclusion** | **10** |  |
| **Reference** | **20** |  |
| **Text Reference** | **10** |  |
| **Clarity** | **10** |  |
| **Formatting**  **Rubric Missing -5%** | **5** |  |
| **Total Points** | **100** |  |