Clustering illusion bias

**Bias Definition**

The clustering illusion is a cognitive bias that involves perceiving patterns or clusters in random or unrelated data. People tend to see patterns where none actually exist, and this can lead to incorrect interpretations of information. This bias arises from the human brain's natural inclination to seek order and make sense of the world, individuals may perceive random events or data points as part of a meaningful pattern or trend. This can occur when there is no actual correlation or connection between the observed elements. The clustering illusion can impact decision-making and judgments, as individuals may base their conclusions on perceived patterns that are not statistically significant or meaningful.

**Ten scenarios of Clustering illusion bias**

1. **Stock Market Trends**: An investor notices that several tech stocks have been performing well over the past week. They attribute this to a new technological advancement but fail to recognize that the trend might simply be a random fluctuation in the market.
2. **Weather Patterns**: A farmer observes that it has rained on every Saturday for the past month. They begin to believe that Saturdays are inherently rainy days, ignoring the natural variability of weather patterns.
3. **Lottery Numbers**: A lottery player notices that the winning numbers seem to have a pattern, such as consecutive or alternating digits. They believe there's a strategy to predicting the numbers when, in fact, each draw is entirely random.
4. **Health Superstitions**: A person starts taking a vitamin supplement and notices they haven't been sick for several weeks. They attribute their improved health solely to the supplement, ignoring other factors like diet, exercise, or simply not being exposed to illnesses.
5. **Gambling Wins**: A gambler wins several rounds of blackjack in a row. They believe they have developed a foolproof strategy for winning, overlooking the role of luck and randomness in the game.
6. **Sports Winning Streaks**: A basketball team wins five games in a row. Fans start attributing their success to a newfound team dynamic or coaching strategy, ignoring the possibility of facing weaker opponents or chance outcomes.
7. **Job Search Success**: A job seeker lands interviews with three different companies in a short period. They attribute their success to a well-crafted resume or exceptional interview skills, overlooking factors like timing, market demand, or sheer luck.
8. **Social Media Engagement**: A content creator notices a spike in likes and shares for their posts over a few days. They believe they've cracked the code for viral content, failing to recognize that engagement rates can vary widely due to algorithms, timing, and audience preferences.
9. **Customer Complaints**: A restaurant owner receives three complaints about the same dish within a week. They assume there's a problem with the recipe or ingredients, overlooking the possibility of random fluctuations or a vocal minority of dissatisfied customers.
10. **Traffic Patterns**: A commuter experiences smooth traffic on their route for several days in a row. They believe they've discovered the perfect time to leave for work, ignoring factors like holidays, accidents, or construction that can disrupt traffic patterns unpredictably.

**User Story: Patterns in lottery numbers**

Sarah, a regular lottery who enjoys testing her luck with weekly ticket purchases. Over the past month, Sarah notices a curious pattern emerging in the winning numbers of the local lottery. She observes that the winning numbers often consist of consecutive digits, such as 12, 13, 14, or alternating pairs like 5 and 7.

Excited by this apparent pattern, Sarah starts to develop a theory that the lottery draws are not entirely random. She begins meticulously analysing past winning numbers, searching for more evidence to support her hypothesis. With each new draw, Sarah eagerly checks the results, hoping to uncover additional patterns or trends.

As the weeks pass, Sarah becomes increasingly convinced that she has cracked the code to winning the lottery. She starts sharing her findings with friends and family, boasting about her newfound strategy for picking winning numbers. Some of her acquaintances start to believe in Sarah's theory, intrigued by the apparent predictability of the lottery draws.

Sarah decides to invest more money into buying lottery tickets, confident that her strategy will lead to big wins. However, despite her efforts, Sarah's luck doesn't seem to improve. She continues to match only a few numbers here and there, never coming close to hitting the jackpot.

Disheartened and perplexed by her lack of success, she decides to seek advice from a statistician friend. With a patient explanation, her friend helps Sarah understand the concept of randomness and the clustering illusion bias. Sarah realizes that the apparent patterns she observed in the lottery numbers were merely coincidences, and the draws are indeed purely random events.

Though initially disappointed by the revelation, Sarah learns an important lesson about probability and cognitive biases. She continues to play the lottery for fun, but now with a more realistic understanding of the odds and a healthy skepticism towards perceived patterns in random data.

We see how the clustering illusion bias can lead individuals to misinterpret randomness and attribute significance to chance occurrences, ultimately influencing their behaviour and decisions.

**Competency questions**

-What captured Sarah’s attention?

Consecutive digits in lottery winning numbers.

-What induced Sarah to believe she had found a winning pattern?

The human tendency to attribute significance to chance occurrences.

**Classes and properties**

**Classes**

**Individual**

Description: a class representing a person that is experiencing the Pareidolia Bias.

**PerceivedPattern**

* + Description: Instances where individuals believe they observe a pattern in data, potentially leading to the clustering illusion bias.

**RandomData**

* + Description: Instances where data or events lack any inherent pattern or order.

**Interpretation**

* + Description: The cognitive process of assigning meaning or significance to perceived patterns, potentially leading to biased conclusions.

**Context**

* + Description: The circumstances or environment in which data or events occur, influencing the perception of patterns.

**Property**

**LeadsToBias**

* + Description: A relationship indicating that the presence of a perceived pattern may lead to the clustering illusion bias.

**BasedOnRandomData**

* + Description: A relationship indicating that the perceived pattern is not based on actual order or correlation but is a result of random data.

**InfluencesInterpretation**

* + Description: A relationship indicating that the context in which data or events occur can influence the interpretation of patterns.

Other properties of the Pareidolia bias have been extracted using chat GPT and then readapted considering the content ODPs in the “Used content ODP section”.

**Key Concepts**

Pattern

Individual

Perception

Bias

Misinterpretation

Illusion

Misconception

Condition

Event

Situation

Activity

**Chosen Framster Frames**

These are the framster frames used for the alignment of the ontology ‘s classes:

**Pattern** (https://w3id.org/framester/data/framestercore/Pattern)

This frame describes the interrelation of a collection of Entities; they may be physical entities or shapes in a recognizable configuration, a pattern of events, or a relation among abstract entities. The pattern is not the individual Entities nor the set of Entities, but an abstraction of their interrelations, as a gestalt. The Cougers are playing in a Wing-T formation tonight. The auditors noticed a suspicious pattern of withdrawals from the maintenance account . The digits of irrational numbers do not repeat in any kind of pattern.

RecognizablePattern(Face)=>fs:Pattern

**PerceptionExperience** (<https://w3id.org/framester/data/framestercore/PerceptionExperience>)

This frame contains perception words whose Perceivers have perceptual experiences that they do not necessarily intend to. For this reason we call the Perceiver role Perceiver\_passive. Comparing the Perception\_experience frame to the Perception\_active frame, we note that for some modalities there are different lexical items in each frame. For instance, whereas Perception\_experience has see, Perception\_active has look at. For other sense modalities, we find the same lexical items in both frames. To illustrate, consider the verb smell where I smell something rotten exemplifies its Perception\_experience use and Smell this to see if it's fresh exemplifies its Perception\_active sense. This frame also includes words which are not specific to any sense modality, including detect, perceive, perception, sense.

Perception=>fs:PerceptionExperience

**ExperiencerObj** (https://w3id.org/framester/data/framestercore/ExperiencerObj)

Some phenomenon (the Stimulus) provokes a particular emotion in an Experiencer. Nightmare on Elm Street scared me silly.

AmbiguousStimulus(Visual Resemblance)=> fs:ExperienerObj

**Entities used from other resources:**

**FOAF**

**Person**: The foaf:Person class represents people. Something is a foaf:Person if it is a person. We don't nitpic about whether they're alive, dead, real, or imaginary. The foaf:Person class is a sub-class of the foaf:Agent class, since all people are considered 'agents' in FOAF.

Participant=>foaf:Person

**Dbpedia**

**Illusion**

An illusion is a distortion of the senses, which can reveal how the mind normally organizes and interprets sensory stimulation. Although illusions distort the human perception of reality, they are generally shared by most people. Illusions may occur with any of the human senses, but visual illusions (optical illusions) are the best-known and understood. The emphasis on visual illusions occurs because vision often dominates the other senses. For example, individuals watching a ventriloquist will perceive the voice is coming from the dummy since they are able to see the dummy mouth the words.

IllusionOfPerception=>dbo:Illusion

**Used Content ODPs**

The following represent the Content Ontology Design Patterns adopted to model the Pareidolia Ontology. Most of these ODP’s classes and properties have been used and combined together during the modeling process.

**ActivitySpecification** (http://ontologydesignpatterns.org/wiki/Submissions:ActivitySpecification)

This work is concerned with supporting a correct and meaningful representation of activities on the Semantic Web, with the potential to support tasks such as activity recognition and reasoning about causation. This requires an ontology capable of more than simply documenting and annotating individual activity occurrences; definitions of activity specifications are required. Current representations of activities in OWL do not meet the basic requirements for activity specifications. Detailed definitions of an activity's preconditions and effects are lacking, in particular with respect to a consideration of change over time. This pattern leverages existing work to fill this void with an ontology design pattern for activity specifications in OWL.

(http://ontologydesignpatterns.org/wiki/Submissions:ActivitySpecification)

**Affected By**

To represent properties/qualities that may affect the status of a feature of interest.

(<http://ontologydesignpatterns.org/wiki/Submissions:AffectedBy>)

**Experience and Observation**

To represent the epistemological "missing link" between a cognitive activity, e.g. the interaction with a cultural object, and any evidence of the effects this activity has on the individuals that are engaged with it; what can collectively be considered as an experience.

(<http://ontologydesignpatterns.org/wiki/Submissions:Experience_%26_Observation>)

**Bibliography**

Wikipedia, *Clustering illusion,* <https://en.wikipedia.org/wiki/Clustering_illusion>

Academy 4sc, *Clustering illusion see the bigger picture*,

<https://academy4sc.org/video/clustering-illusion-see-the-bigger-picture/#:~:text=The%20clustering%20illusion%20occurs%20when,across%20an%20entire%20data%20set>.