

# MORAL ALIGNMENT

## 1. Definitions

### 1.1 Axis:

- **Impulse**: Personal tendency towards a category of act.
- **Structure**: Type of response/action to **tendency**.

### 1.2 Impulses:

- **Good**: Tendency to filter decisions based on inner believed **good** values.
- **Evil**: Tendency to filter decisions based on personally believed **evil** values.
- **Neutral**: **Not** having any feeling of good or evil, putting **survival & situation diffusion** above.

### 1.3 Structure:

- **Lawful**: Tendency for **sticking to rules/laws** of a system in response to **impulse**.
- **Chaotic**: Tendency to deliberately **break rules/laws** of a system in response to **impulse**.
- **Neutral**: **Not** having any feeling towards or against laws/rules, thus responding to **impulse** with most efficient option.

## 2. Facts & Myth Busters

- Understanding of objective truth and reality doesn't make someone naturally **neutral**, but inner tendency does.
- Both the axis are result of **inner tendency** or **feeling**, not visible actions.
- Clashes among one or both same axis are expected for perceived difference in values.
- Being aligned with **neutrality** on any or both axis doesn't equate to stable behavior.
- While everyone has some of the each tendency, one dominates and controls the individual.
- By doing something that doesn't align with inner tendency on any axis, it produces a feeling of unease in any form.
- **Evil** impulse doesn't only mean making effort to cause trouble for others, but could also be to pursue or respond to something unnecessary, costing others negatively, over survival (**neutral**).
- Taking hard decisions can pervade the basis of moral alignment.

## 3. Grid Of Combinations

### 3.1 Combinations:

Impulse/ Structure	Good	Neutral	Evil
Lawful	Lawful Good	Lawful Neutral	Lawful Evil
Neutral	Neutral Good	True Neutral	Neutral Evil
Chaotic	Chaotic Good	Chaotic Neutral	Chaotic Evil

### 3.2 Mental Processing:

Impulse/ Structure	Good	Neutral	Evil
Lawful	When faced with choices under favorable conditions, chooses what seems to align the most with perceived <b>good values</b> by <b>strictly adhering to rules/laws</b> of a system.	When faced with choices under favorable conditions, chooses what seems best to <b>survive</b> in the given situation or <b>settle it</b> , but by <b>strictly adhering to rules/laws</b> of a system.	When faced with choices under favorable conditions, chooses what seems to align the most with perceived <b>evil values</b> by <b>strictly adhering to rules/laws</b> of a system.
Neutral	When faced with choices under favorable conditions, chooses what seems to align the most with perceived <b>good values</b> by choosing the most <b>optimal option</b> for responding to impulse.	When faced with choices under favorable conditions, chooses what seems best to <b>survive</b> in the given situation or <b>settle it</b> , but by choosing the most <b>optimal option</b> for responding to impulse.	When faced with choices under favorable conditions, chooses what seems to align the most with perceived <b>evil values</b> by choosing the most <b>optimal option</b> for responding to impulse.
Chaotic	When faced with choices under favorable conditions, chooses what seems to align the most with perceived <b>good values</b> by <b>deliberately breaking rules/laws</b> of a system.	When faced with choices under favorable conditions, chooses what seems best to <b>survive</b> in the given situation or <b>settle it</b> , but by <b>deliberately breaking rules/laws</b> of a system.	When faced with choices under favorable conditions, chooses what seems to align the most with perceived <b>evil values</b> by <b>deliberately breaking rules/laws</b> of a system.

## 4. Procedure (v1.0.0)

1. Understand about the axis and terms from **\*section 1** and **2**.
2. From the **section 1.2**, rate yourself for each **impulse** on a scale of [1, 5].
3. From the **section 1.3**, rate yourself for each **structure** on a scale of [1, 5].

4. Add up the results from **step 1,2** to make the  $3 \times 3$  grid.
  5. Choose the combination with highest score.
  6. Confirm the result by reading same combination from ***section 3.2***.
-