## **Anti-Misalignment**

#### **Richard Kim**

Member of Technical Staff @ Anthropic Prev: Staff Engineer @ Airbnb



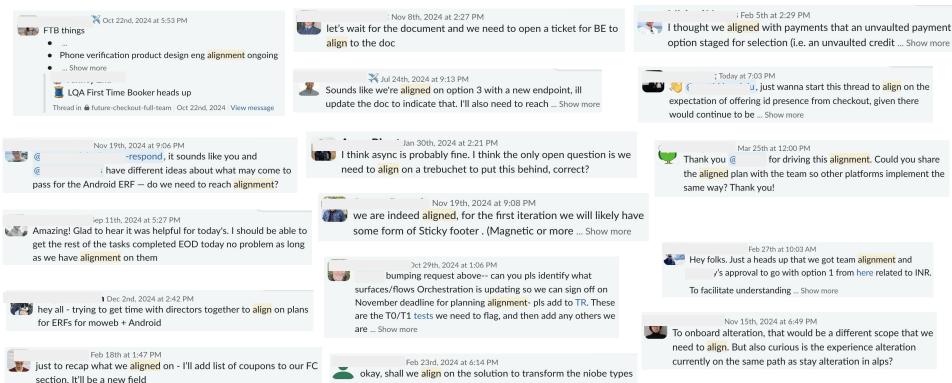
## "Alignment"

Goal: You will feel weird using this word after today

Raise your hand if this is something you worked towards or cared about in the last 6 months

#### Anyone who didn't raise their hand is a liar

cache? @



and pass through the passport modules instead of relying on the

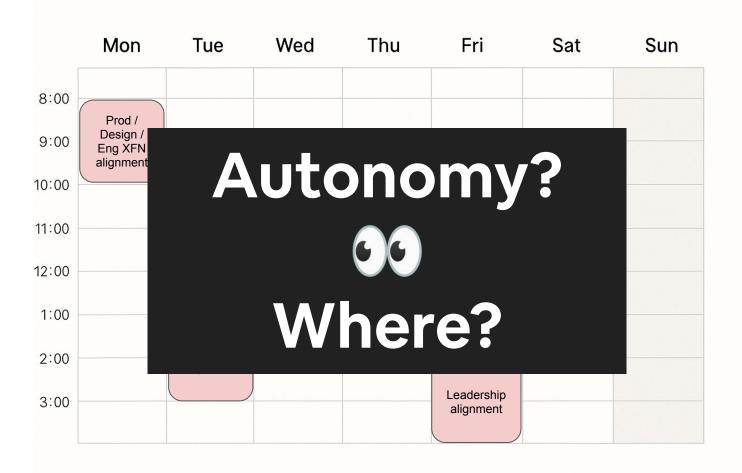
#### Some common ones

- Let's make sure we're aligned with Design on this
- The goal of this meeting is to get alignment on our launch readiness
- Is OtherTeam aligned on the new requirement?
- Can we align on how we manage bug intakes?
- Sharing alignment on XYZ
- Here's the spec we aligned on

## "Alignment"

Who can tell me what it is

### **Autonomy?**



#### **Examples - not about autonomy**

- The goal of this meeting is to get alignment on our launch readiness.
  - The goal of this meeting is to **agree** if we are ready to launch
- Is OtherTeam aligned on the new requirement?
  - Object of the Does Other Team know about the new requirement?
- Can we align on how we manage bug intakes?
  - Can we agree on a process to manage bug intakes?
- Sharing alignment on XYZ
  - Sharing our decision to XYZ
- Here's the spec we aligned on
  - Here is the spec we agreed to

# "Alignment"

WHAT DOES IT MEAN???

#### Why do I care?

- Obsessed with this question for more than a year now
- In many ways, I owe my career to "alignment work"
- The work is hard, it's how we find gaps, it's how big projects move forward. It's how we develop strong relationships across ownership lines



## "Alignment is coercion"

Annie Duke on "Lenny's podcast"

#### Annie Duke - "Alignment is coercion"

Proceeded to NOT EXPLAIN HERSELF AT ALL

Shook me to my core, has all this hard work been BS? It sometimes feels like BS

Oh actually, I have no idea what alignment is

Over the past year I asked many ICs, asked managers at other companies, asked staff+ engineers, mentors outside of airbnb and...



# Nobody knows what "alignment" is

(From my conversation with many folks)

Sure, people gave definitions, but actions did not line up with those definitions

#### People could not shut up about misalignment

- Strong feelings
- Pain
- Frustration
- Example after example after example

#### **Thesis**

Nobody knows what alignment is

But everyone knows what misalignment **feels** like.

## "Alignment"

An exercise

#### Align with me

- Alignment is not really a thing that you do
- Alignment is a state that evokes a feeling

#### Alignment is a **flow state** between people

- Flow state Mihaly Csikszentmihalyi
- Group flow state The group feels like a trusted extension of yourself. You feel fully utilized to focus, make decisions, and move.
- What do we think? /

#### How do you get into individual flow?

Ask for the room: How do you get into flow at work?

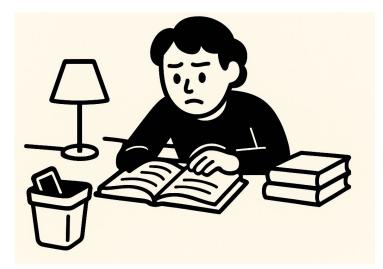
# How do you get into flow? - Common example

This 👉

Is not actually how to get into flow, but how to prevent falling out of flow

Losing flow is painful, and it's what you remember

"I woke up early, brought my work to the library, turned off my notifications, quit all my apps, deleted Instagram, threw my phone in the trash, turned on white noise, and just stared at it until it started happening"









#### Breaking flow is painful

- Breaking flow is so painful that we think about flow in terms of not breaking it
  - "Turn off my notifications"
- Misalignment is so painful that we think about alignment in terms of preventing misalignment
  - "Is OtherTeam aligned that XYZ?"
    - -> "Are we sure enough context is shared between us that this won't bite us later?"

#### **Updated thesis**

Nobody knows what alignment is, but everyone knows what misalignment feels like.

Thus, most "alignment" is just **preventing misalignment** 

We'll call this "anti-misalignment"

#### Anti-misalignment (n)

1. The act of clamping down on misalignment as a brute force strategy to achieve alignment.

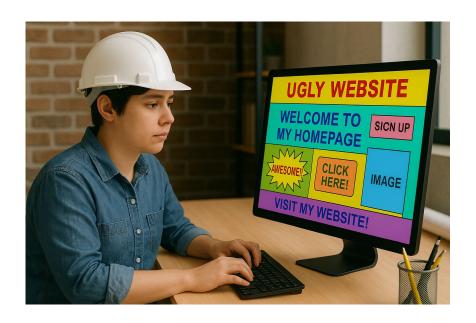
Synonyms: Lazy alignment, reactive alignment.

#### People could not shut up about misalignment

From these conversations, I found that Misalignment is **felt** whenever:

You discover a decision that feels both surprising and harmful to your goals

#### Misalignment requires **Discovery**





No discovery -> no misalignment

#### Misalignment requires **Surprise**







#### Misalignment requires <u>Harm</u>





No harm -> no misalignment

#### Sor

1.



#### Some

- 1. **Rec**
- 2. **Ma**l



Sc 0,1% 2. 3.

#### Some funny ways to reduce misalignment

- 1. Reduce the discoverability of decisions.
- 2. Make decisions less surprising.
- 3. Make decisions less harmful to your goals.
- 4. **Stop making decisions**. If nobody makes any decisions, then nobody can feel misaligned!

#### If any of these felt relatable in any way

You have fallen victim to "Anti-misalignment"

#### Update the thesis

Nobody knows what alignment is, but everyone knows what misalignment feels like.

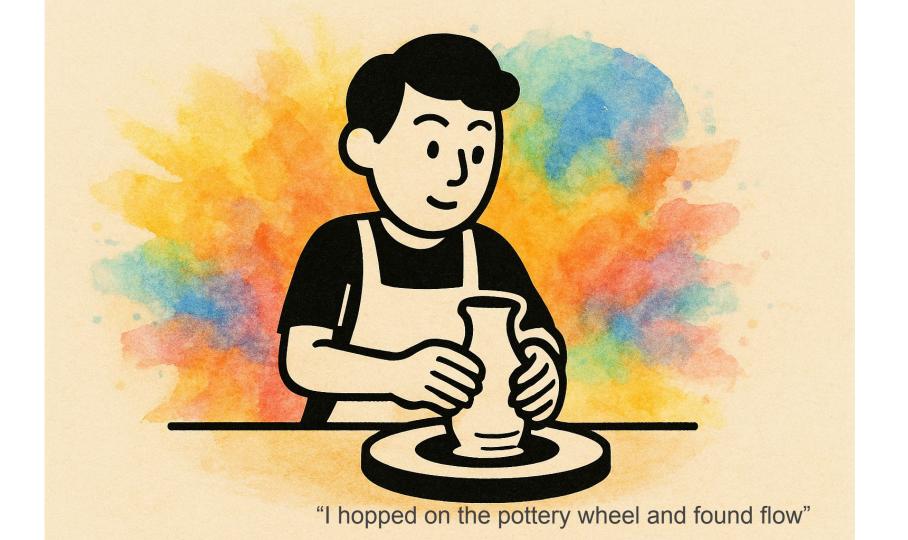
Thus, most "alignment" is just preventing misalignment ("anti-misalignment")

"Anti-misalignment" is less effective than true, proactive alignment

## Return to flow example

"I woke up early, brought my work to the library, turned off my notifications, quit all my apps, deleted Instagram, threw my phone in the trash, turned on white noise, and just stared at it until it started happening"





#### Weak knockoff flow



"I woke up early, brought my work to the library, turned off my notifications, quit all my apps, deleted Instagram, threw my phone in the trash, turned on white noise, and just stared at it until it started happening"



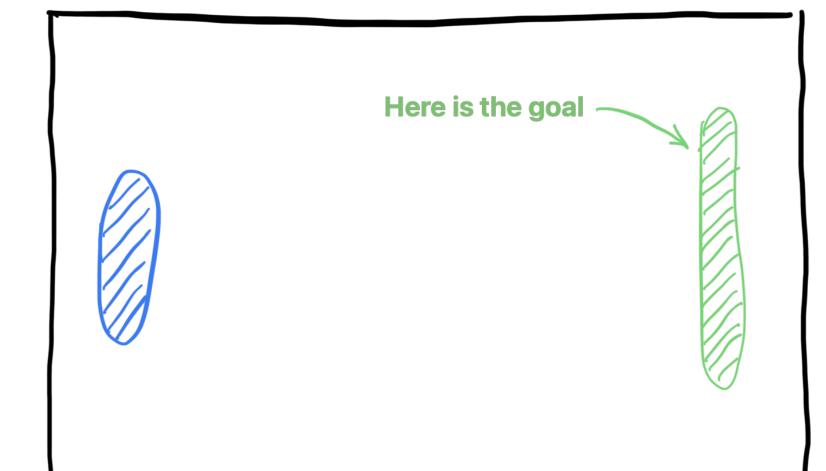
#### Same for alignment

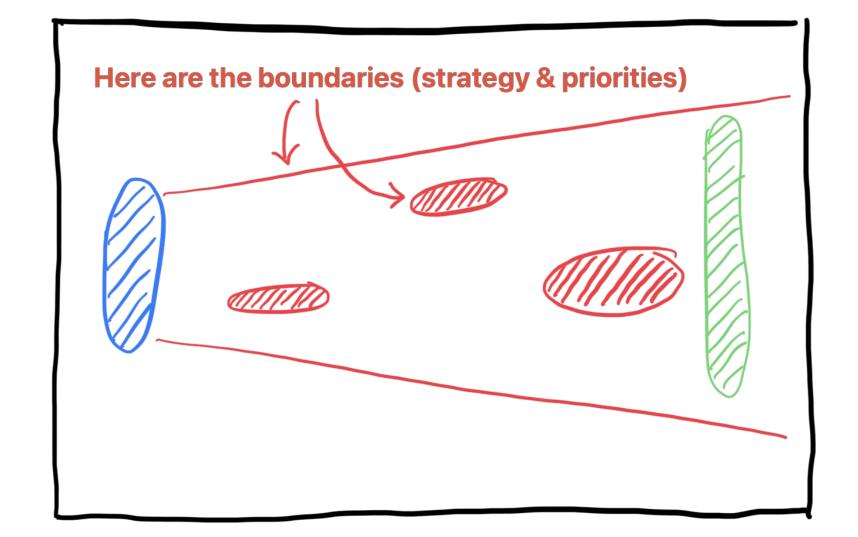
- Focusing on stopping distractions is a worse way to get to flow
- Focusing on **stopping misalignment** is a worse way to get to alignment

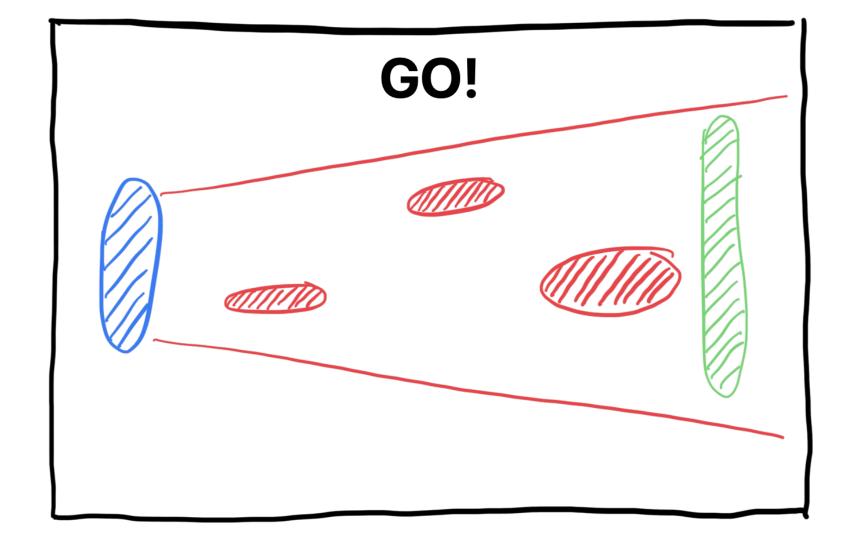
Let's look at some examples of alignment vs anti-misalignment

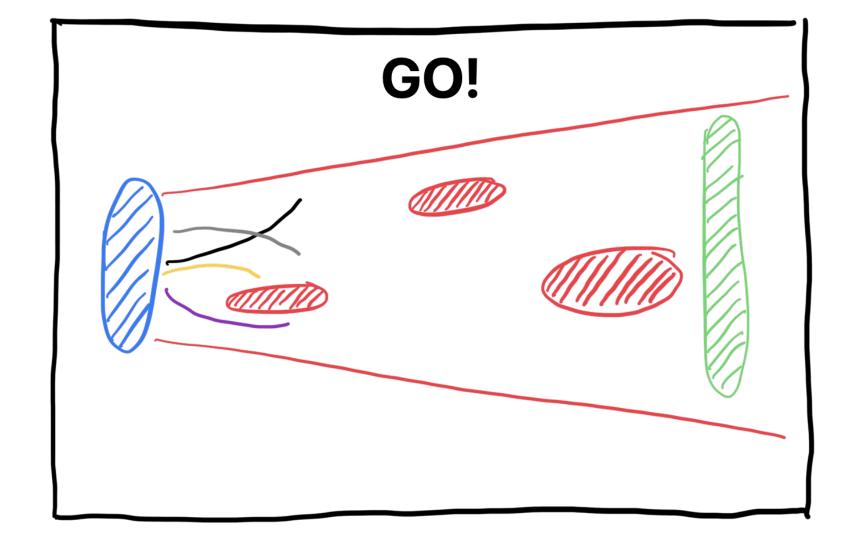
## Alignment example

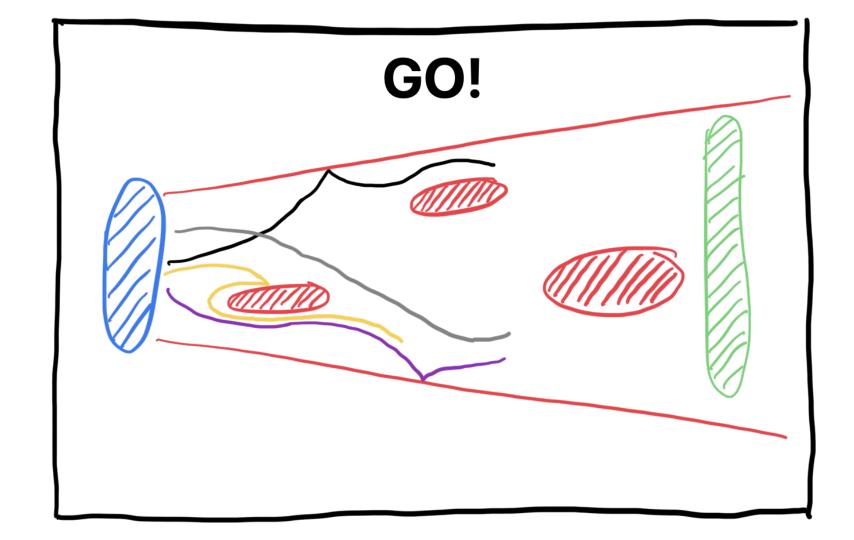


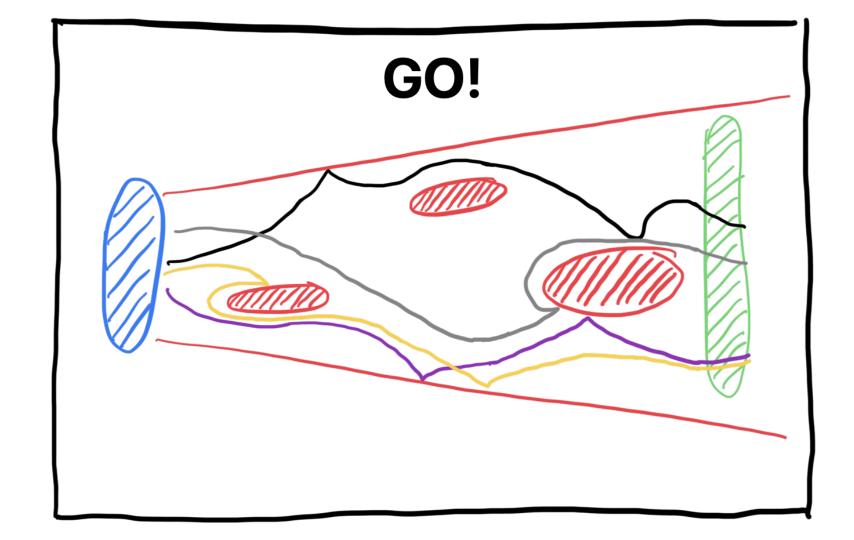


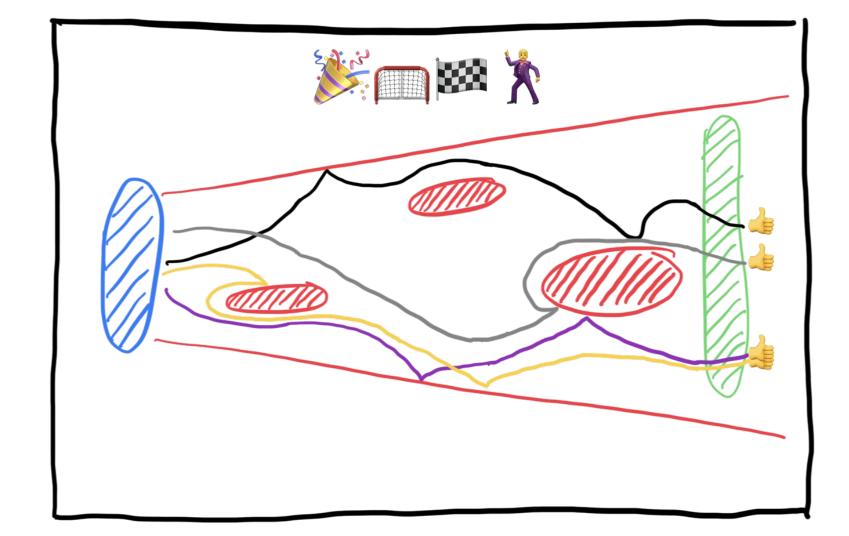




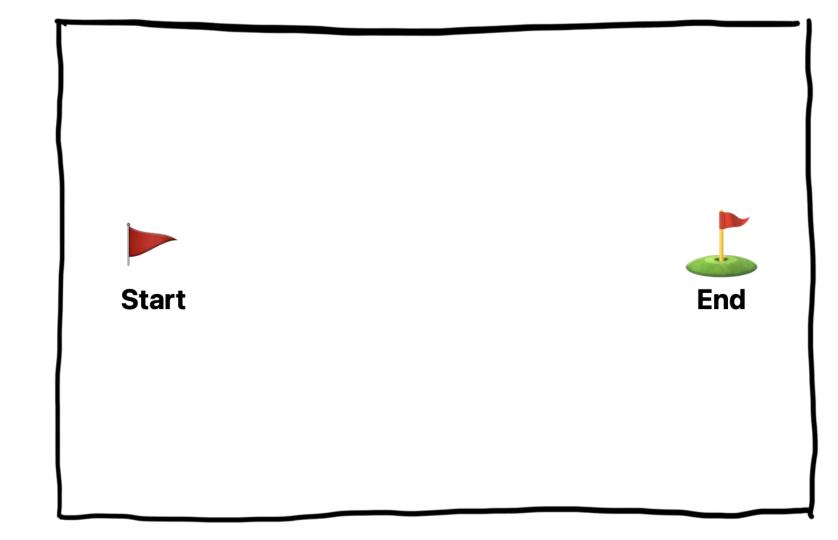


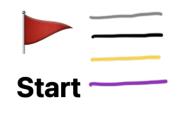




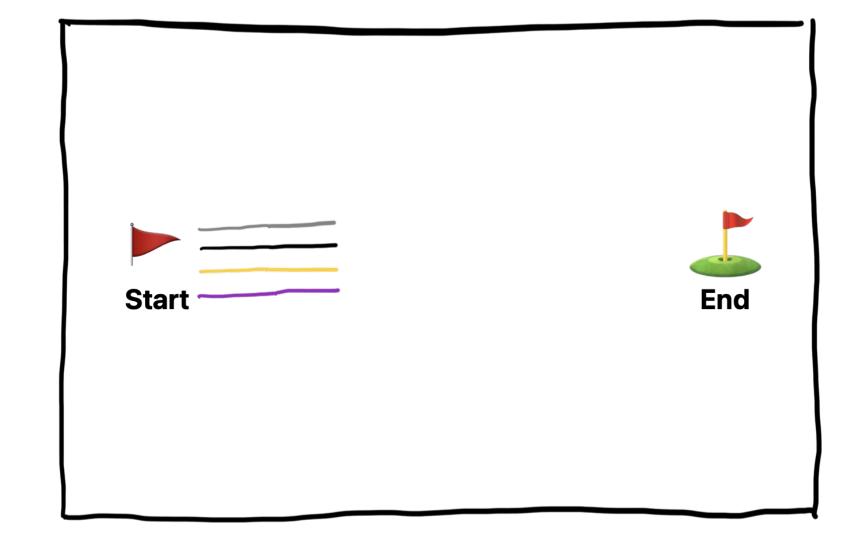


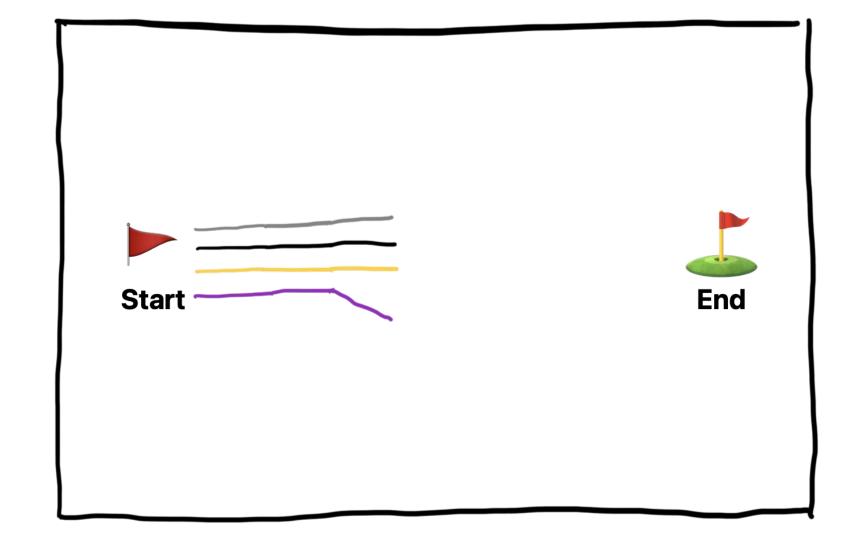
Anti-misalignment example

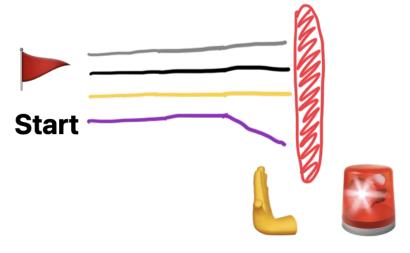




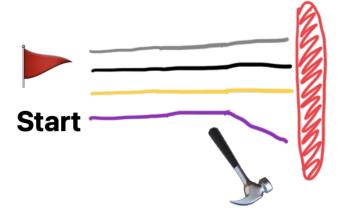




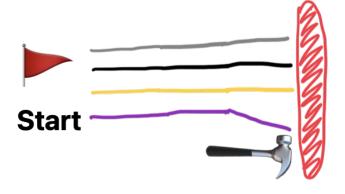




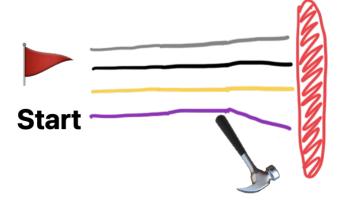








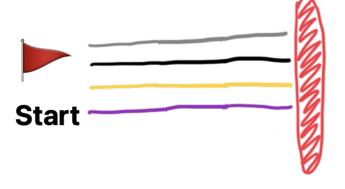














## Misalignment clear







#### **Anti-misalignment**

Anti-misalignment doesn't aim for clarity - it aims for control.

Since no one understands the **strategy** well enough to act independently, the only way to stay "aligned" is to move in lockstep.

**Every decision is a potential landmine**, so no decisions can be made independently

#### What do we actually do about it?

"There are only two hard things in Computer Science: cache invalidation and naming things" - Phil Karlton

- Police terminology?
  - o Doesn't work, I tried
- "Alignment" means everything and nothing all at once
- If you try to convince someone they are not aligned, they will convince you that actually, they are (for fear of another meeting)



## Strategies

#### 1 - Just stop using the word

- Try anything else
- Context, goal, strategy, priority, motivation, approach, agreement
- "Higher level"
- Shake out a different answer

Whether we have "alignment" is not actually the goal because remember, nobody knows what "alignment" is.

Instead, make sure people have **clarity**. Make sure they understand the **strategy**. Make sure they know our **priorities** and the **context** behind the decisions we make.

The goal isn't "alignment", then the goal is **progress**, **flow**, **movement**, **autonomy**.

#### 2 - Treat alignment as variable

- Alignment is not static
- Earlier animation example isn't perfectly accurate
- New requirements, new priorities, new people = new alignment

Instead of aiming for one perfect alignment artifact at the start, focus on enabling group flow right now, and be ready to adjust as things evolve.

#### 3 - Advocate for autonomy

- There's a difference between alignment and agreement
- Agreement is good, but consider looking for the pattern or the strategy to how the agreement was made. Can it be generalized?
- How can you prevent the next alignment meeting?

#### Example of all 3 (different terms, variable, autonomy)

**Meeting host**: the goal is to align on what to do about issue XYZ that came from our project colliding with a new company initiative.

**Group**: (discusses, arrives at a decision)

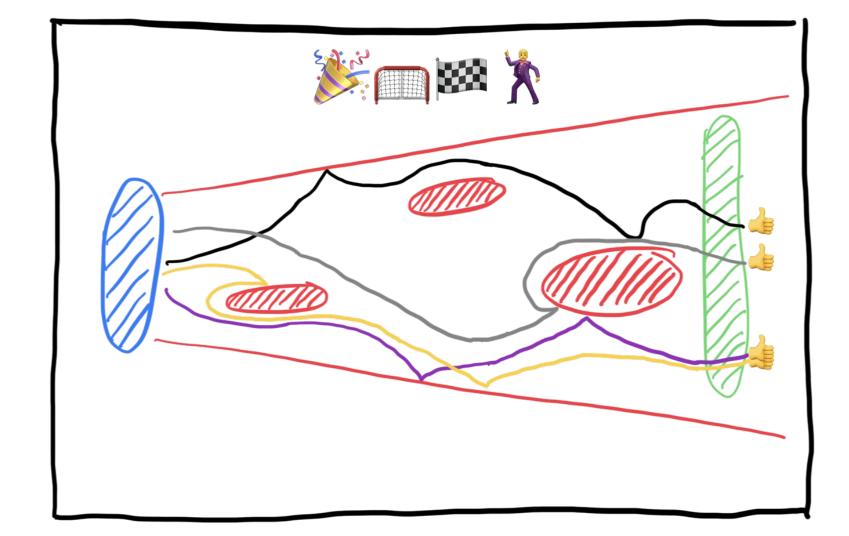
Meeting host: Great! We have alignment

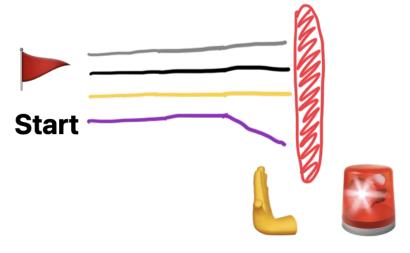
**You**: Sounds like we made this decision because **Y priority now outweighs Z**. Do you think that **tradeoff holds more broadly**? If so, we could share it with our partner teams to **help them make similar decisions** without needing to sync every time.

Good job using alternatives to "alignment", remembering alignment is variable, and advocating for autonomy!

## Conclusion

# "Alignment"







## Are we all aligned?

Great! Let's get lunch

## Examples

We know from past projects that this page is really sensitive. Small changes often lead to really big results (both good and bad). So, we prioritize testing assumptions iteratively over launching big bundles of features all together.

This is outlining a priority and the resulting execution strategy. This enables alignment and autonomous decision-making. It is the "why" behind the decision, which helps empower others to make decisions

## Let's align on the process of communication between our two teams for how we will operate.

- This is an agreement, not a strategy. This doesn't help determine how to make changes in the process in the future, just agreeing on a decision right now.
- Agreement isn't a bad thing and is probably what is necessary in many cases for speed. The purpose of drawing the distinction is not to discredit "agreement", but to encourage people to actually seek alignment after the agreement is made.

How this could be alignment: "Let's align on what the goals are for our teams and derive a process that matches those goals. Hopefully we can preempt some of the gaps that emerge, but if not, let's discuss how we will adjust and improve our process over time".

• You can see that this takes more depth, so a short agreement could be appropriate depending on the context.

We are in a phase of discovery to find our next big thing. We should use data, research, prototypes, and quick experiments to try every idea and see what sticks. We should not be afraid to take a slight brand hit by launching something unpolished. We also should not be afraid to end an experiment that is working, but won't be a huge opportunity

This is also a high level strategy. It provides clear priorities and tradeoffs to give a wide breadth of autonomous decision making.

Here is what we are launching next quarter! Notice that XYZ is not in the build, but ABC is in the build. The reason is 123, but we're happy with that trade off. If we're in alignment, let's lock it in and finalize the roadmap!

This is an agreement on what we are building, it doesn't enable individuals to make tradeoffs and decisions. What happens if an issue is discovered and another feature can't make the build? We would have to "align" again.

It is close though! If you add "we're happy with that trade off because priority X takes precedence over priority Y for our goal to Z", then that could enable future decision-making.