NoEstimates Game Facilitator Guide

Last update: Apr. 22, 2019
Explore sources of variation and alternatives in forecasting
Contact: Matt Philip (@mattphilip, matthewrphilip@gmail.com)
For updates, please visit http://noestimat.es

Objective

Learn what and how much different factors influence delivery time and understand options for forecasting

- Total time: 1.5 to 3 hours (includes debrief)
- Teams: 1-8 teams
- Players: 4-50+ people (4-8 players per team)
- Inspired by Vasco Duarte's NoEstimates book and Russ Healy's GetKanban Game

Game Contents (per team)

- Team name placard
- Game board
- Work cards (face down, in the Options area)
- Event cards
- Role name tags (one per person)

- Role cubes (four per person)
- Scorecard
- Scatterplot chart
- Cumulative-flow diagram
- Six-sided die
- Eight-sided die
- Day-indicator token
- Dry-erase marker
- Blocker spinner
- Mini Post-it notes (for blockers)
- Team-dependency cards (remove the team's own color and shuffle into others so that each team's stack should not have its own color)

Overview

Spend about 10-15 minutes walking through the slides to explain the rules (following). Give the teams a solid 60 minutes of game play (or enough for at least one team to complete at least Day 10 and 10 cards). Then spend 15-30 minutes debriefing. Each team has more cards than they can complete, so the facilitator should stop the game based on time remaining for the debrief rather than have the teams complete a certain number of days or cards. The events of the game, from play through debrief, take players through a progression of estimating approaches, from deterministic estimates based on guesswork, to deterministic forecasts based on data, to probabilistic forecasts based on data — with your help as facilitator!

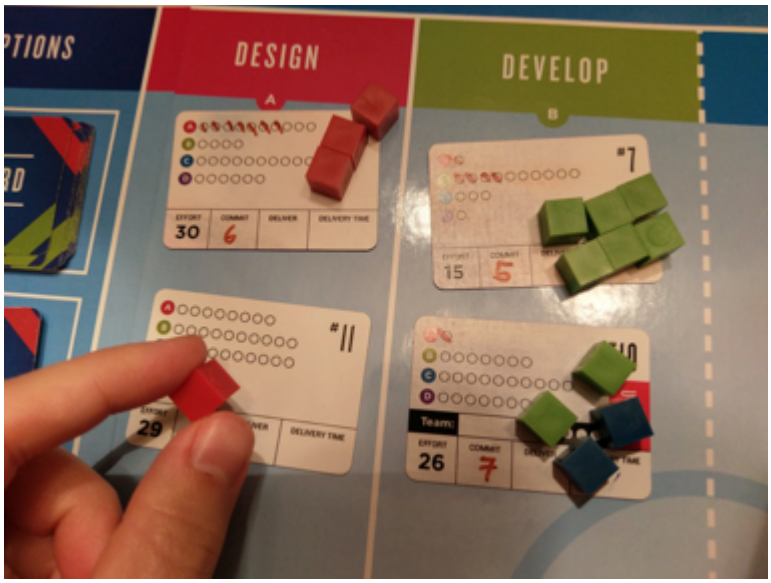
Facilitator Script

1. Welcome everyone to the game. You might say something like "Who has ever had to answer the question 'when will this be done?'? We'll explore different options for answering that question, along with lots of other questions related to delivery in knowledge work."
2. Walk through the presentation slides (https://mattphilip.wordpress.com/noestimates-game/ and at https://github.com/pip72/no-estimates-game).
3. When you ask the group to decide teams and roles, tell them that their signal to you that they are ready will be when each player is wearing a role name tag (Designer, Developer, Tester, Deployer)
4. Allow teams to be non-uniform, comprised of different numbers of people or even a team with all the same specialty. It's even acceptable to not staff one of the roles, since other roles can work outside of their specialty.
5. You may optionally share the following work notes, which may influence their decisions (e.g., they might staff fewer or no designers because design doesn't have

as much total work):

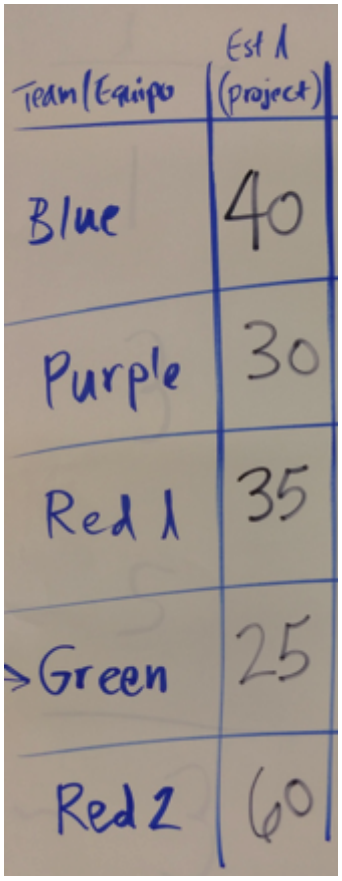
a. Only 75% of cards have Design work
b. Dev and Test together accounts for ~ 2/3 of total effort

6. Tell them that the plastic cubes represent their daily effort (we work at a sustainable pace!), and that each person should take four cubes of the color corresponding to his or her role specialty. They then place those cubes on cards to help them during their daily planning meeting so that they don't have to keep track in their heads how much effort they've "spent" each day.



7. Before starting, give the teams a few minutes to provide an up-front estimate on how many days it will take them to complete the project. Write them on a white board or easel under a column labeled "Initial Estimate" (create two additional columns but leave them blank for now).You'll probably have a range of estimates, which provides a point of meaningful humor, if you observe aloud that "It's interesting that each team has the exact same backlog — with a completely known amount of effort! — and still you gave me a wide range of estimates! What's going on!?" Someone might point out that the team sizes may differ. You might concede that point, but question whether it should make as much difference as some of the estimates indicate. Or at least point out that teams with the same number of people

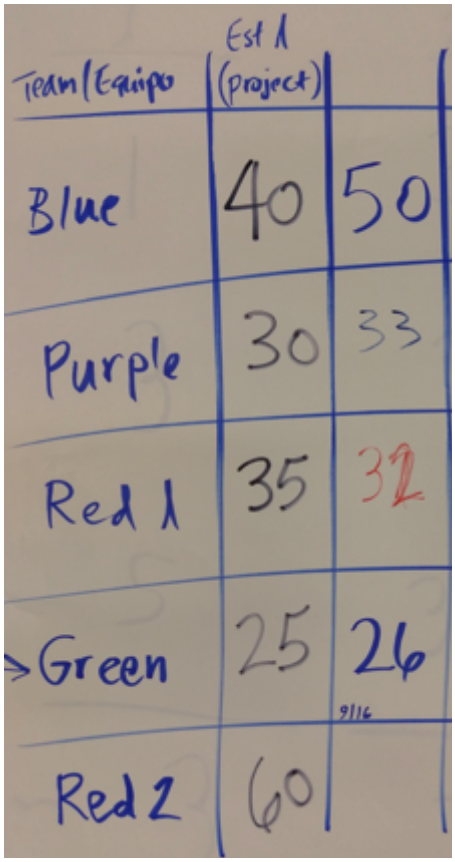
shouldn't vary.



Team/Equipo	Est 1 (project)
Blue	40
Purple	30
Red 1	35
Green	25
Red 2	60

8. Teams can now begin play!
9. At the end of Day 10, teams draw an event card that has "Vince from the PMO" demand that they re-estimate. They'll understandably be disgruntled, and some players may even be tempted to simply repeat their original estimate. But encourage them to take Vince's question seriously and give an honest answer, which

you (or they) should capture on the board in a second column:

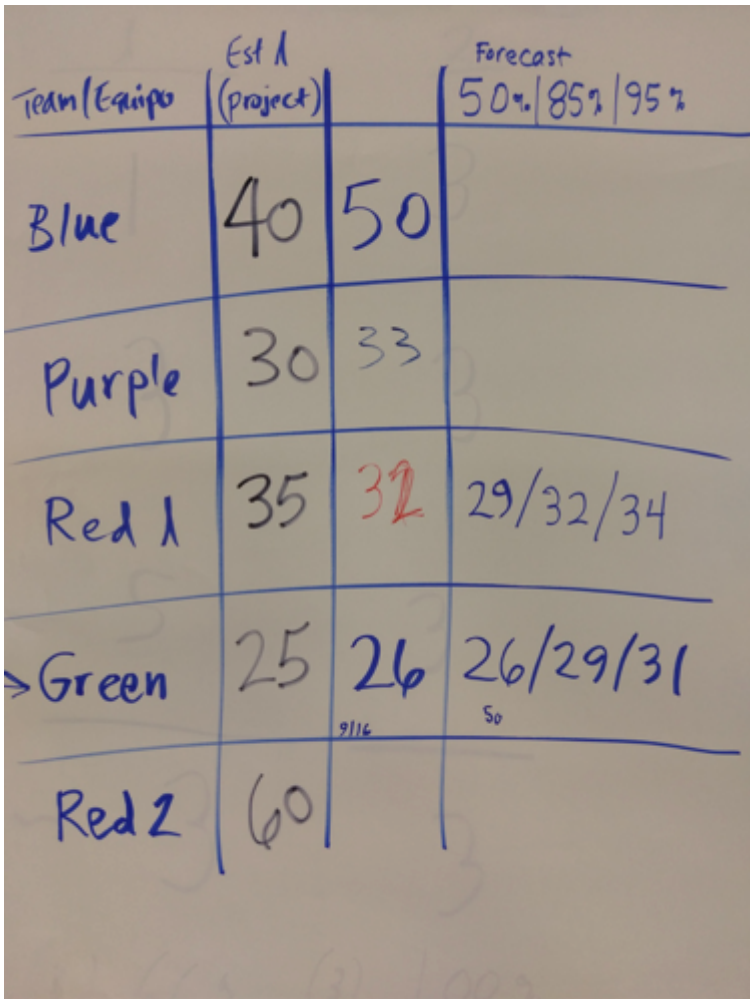


Ending the Game

1. Depending on how much time you have, you can let teams play as many days as they can, allowing at least 15-30 minutes for the debrief.
2. At minimum, aim to have at least one team complete 10 cards (it doesn't need to be card numbers 1-10). This way, you'll have enough data to do a probabilistic forecast (e.g., the Actionable Agile requires at least 10 data points in order to forecast). Also, try to have as many teams as possible reach Day 10, which is when they draw an event card that has "Vince from the PMO" require them to re-estimate.
3. At the end of each team's day 10, be sure to visit the team (prompted by the Event card). The event card requires them to provide a new estimate of the entire project. The facilitator should also at this time create a probabilistic forecast (using their delivery-time data), which they will be able to compare with their estimate.
4. It's not necessary for all teams to remain on the same day cadence, but they need to maintain the integrity of each day's work allocation (e.g., one worker can't spend more than his or her daily allocation) – that's why the colored cubes are helpful. Remind them of the context-switching penalty for working in two teams in the single day, as some players will have one cube left at the end of their own team's day and want to

give it to another team but will forget to incur the penalty, which would take away that cube.

5. As teams start to complete stories, go around to each team and record in a spreadsheet their delivery data:
    - Start date
    - Delivery date
    - Estimate
    - Card #
6. When you're reading to begin the debrief, enter the data into your favorite probabilistic forecasting tool (recommend Troy Magennis's tool or Dan Vacanti's Actionable Agile tool).



Debrief and Discussion Questions

Begin the debrief by reviewing the three-stage progression of thinking behind their estimates. For example:

"At the beginning of the game, I asked you to estimate how many days it would take you to complete the entire project. Can someone briefly explain how you came up with your initial estimate? What factored into it?"

- Most teams will say that they took the total effort data (average of 21 per card times 25 cards = 525) and calculated a number based on their capability (eight players each with four effort per day = 32 effort/day) and some fuzzy contingency or "gut" feeling. You might note those factors ("Effort + Gut Feeling" at the top of the first column). You might also ask if anyone used Story Points and Velocity — why or why not? Occasionally, a team will report that they assumed their estimate needed to be 30 days or less. Ask why, and they'll usually reply that the number of days on the board went up to only 30. This is a perfect time to introduce one of many human biases in estimating, that of anchoring. ("Have you ever had a manager or someone else say 'Please estimate this -- I think it's a six-month project, though'?")

  Then ask someone to explain how the team answered Vince (Day 10 event; their second estimate). Did their approach change? Usually, teams will look at their finished work in order to forecast. For instance, they'll see that they've completed 10 cards in 10 days an extrapolate that for 25 cards and estimate that they'll finish on day 25. This is still a deterministic forecast based on average, but it's an evolutionary step from the first estimate because it's using real delivery data. You can use this opportunity to talk about the "flaw of averages" and how we still need to accommodate more than one possible outcome, rather than a single, deterministic outcome.

  Then you move to the third "estimate," which you will perform for them using a Monte Carlo simulation. Here, you will show a range of possible outcomes, or a probabilistic forecast. I usually show 50th percentile, 85th and 95th. Depending on how "predictable" the teams were in their delivery (you can refer to their scatterplot charts), this range will vary (some will be only a couple of days apart, while some may be many days).

  Other debrief questions:

- In this game, you had "perfect estimates" of effort; that is, the cards told you exactly how much effort each required. And still, you were not able to estimate very well up front. Why not?
- What sources of variation did you experience? Some that people may have experienced in the game:
  - Context switching (Duarte calls this "focus factor")
  - Work in progress (strong lead indicator)
  - Team dependencies
  - Team composition

    - Availability of specialists
    - Rework
    - Blockers
    - Selection policy (urgent cards “jump” over other WIP, incurring flow debt)
    - Time spent estimating! (After Day 10, when Vince requires it)
- When we estimate, what is the underlying assumption about its relationship to delivery time?
- What was the correlation of delivery times and estimates? Show correlation numbers from spreadsheet (perfect correlation is 1 or -1; good correlation is at least .5 or -.5; expect low correlation, such as .4 or less)
- What was the best strategy? What tradeoffs did you make?
- How do we reduce variation? Reduce variability by leveling out demand (e.g., small user stories), heijunka
- Who followed a policy (formal or informal) of maximizing utilization? How did that work out for you? Or did you choose to limit WIP and have less overall utilization? (Notice whether any team had a smaller average WIP than number of players – “Could you do this in real life?”)
- Did you choose to invest in automated deployments? What was the effect? In real life, how would that impact predictability?
- How did you calculate your original estimate? Your MVP?
- How did you estimating/forecasting approach change? (Usually, teams in the game follow the pattern of
    - 1. Estimate based on effort
    - 2. Estimate using data (they look only at how many cards they’ve delivered and extrapolate from that). Discuss how this approach, while an improvement insofar as it uses data, still relies on an average. Discuss the flaw of averages (Team A delivers stories in 1, 3 and 5 days, while Team B delivers in 3, 3 and 3. Both average 3 days but one has a 66% chance of delivering in 3 days, while the other has 100% chance.)
    - 3. Probabilistic forecast: Use date to create a probability and a range.
- Do we have better ways of forecasting?
- If estimates are unreliable, what options do we have?
- Show probabilistic forecasts of the teams’ data (use Actionable Agile or other tool).
- How did your daily planning meetings feel? How did that differ from your real-life standup meetings? Why?
- Did anyone end a day with an unused cube? Why? How did that feel? What was the tradeoff you were making?
- In your first estimate, some of you may have taken an “underpromise, overdeliver” attitude. Why do we do that? Why do we incorporate a “buffer” or “contingency” into our estimates? When we forecast with data, it can take the pressure off the us-against-them relationship, because it simply shows us how our system works, without the mistrust of buffers and accusing teams of padding their estimates.

- Did any of you change a policy mid-way through the game? How did you know whether it had its desired effect? Can you point to any of your charts to explain?
- Did the fact that you all "work for the same company" change your behavior with respect to optimizing for your own team? In what ways might you have optimized for the entire company? ("Systems thinking")

More Discussion of Sources of Variation

Time spent estimating

From Dan Vacanti's "When Will It Be Done?" book:
"Strange as it may sound, estimating and planning usually serve to make you less predictable, not more. Think about it this way: during all of the time that you spend on estimation and planning, what are you not doing? Working. That time could otherwise be spent delivering customer value. The more time you spend planning and estimating the less value you are able to deliver. In fact, if you interrupt work to do estimation, as so many companies do, estimation is actually making you less predictable. Think about how many times you are interrupted for an estimate and how much time that costs. How much value did you get for that cost?"
This is essentially a form of the agile manifesto value of "Working Software over comprehensive documentation" — estimating is a form of comprehensive documentation.

FAQs

Can we pick the cards to play?
No. Assume that the product owner has prioritized these in the correct order. Teams only see the face of the card when they decide to pull (commit to) it, though they can pull as many as they like. (Though important, work selection, sequencing and scheduling is not an intended learning in this game.)

If our team doesn't have an actual person playing a particular role, do we get to use an "imaginary player" for that role and get four effort points to use?
No. Teams that do not have a role represented do not get to play with an "invisible" player!

If card is blocked, can you get another team's help on it or does it need to be unblocked first?
If a card is blocked, you cannot do any work on it (by anyone) until it is unblocked.

When the team is locked in the room to estimate and not doing work (end of Day 10), do they spin to block/unblock?
No, the team doesn't do any work that day, including spinning to unblock.

Do you actually use the $$$ for anything? Or just to get them to hurry up urgent tickets?
It's used to keep score of value delivered. But it's really meant as a way to provoke selection-policy decisions. In the debrief, we rarely discuss the scores but do talk about the impact of selection policy as a source of variation.

Does the player from another team have to be a certain role to help with the dependency work?
No. Help from another team to complete the team-dependency work can be from any role.

Optional Rules and Add-ons

Single-Board Play

If you have only one board (or two boards), you can simulate team dependencies. Whenever the team draws a work card with a team dependency, you as facilitator simply simulate that team's contribution by rolling an eight-sided die:

- 1-4: Tell the team that the dependent team gives them this number of effort cubes for the day.
- 5-8: Tell the team that the dependent team can't help them today and to check back tomorrow.

Depending on the team's real-life circumstance, you might choose to further recreate a similar scenario, such as a dependent team in another time zone who can't respond to their request for at least a day.

Use the following rules if you have extra time or would like to introduce additional learning.

Customize the Probability of Events

At setup, ask the individual teams the following:

- How often do you have rework?
- How often do you have a blocker?

Depending on their answers, you can customize the game to reflect their real-life probability of events. For instance, the standard game rules have a 1 in 4 (25%) chance of rework and blockers occurring. But if a team reports that they have rework only 10% of the time, you might use a 10-sided die, with 1 being the "hit." Or if a team gets blockers 33% of the time in real life, change the hit rate to a roll of a 1 or a 2 on a six-sided die.

Blocker Clustering

- Keep track of your blockers

- For each day that it is blocked, dot the card
- At the end of the game, count the blockers:
- Number/rate of occurrence
- Total number of days blocked
- Reason for block