선도인재양성 중급
(5주차)

AI + X
Advanced Project

✓ Schedule

- **5주차(3/31)**: AI 서비스 기획을 위한 **리서치**
- **8주차(4/21)**: 중간 프로젝트 발표 (평가 대상, 20%)

1. 확정 컨셉

2. 초기 POC

✓ Today's Agenda

- Project Progress Report and F/B (1H)
- User Research for AI Services (1H)
- Group Discussion (1H)

Class Goals

AI 서비스 기획 및 POC 개발

One “Big” Project

cf) three projects

Impact!

Analysis

Project Brief

“Develop a Conversational Experience based on Generative AI” (Chat GPT API)

Midterm
Grading Rubrics
평가 지표

✓ Midterm Presentation

Mainly about three things

1) Research findings and insights 사용자 리서치를 통한 Insight 도출

translated into

2) Design question —— 문제(사용자 니즈, Painpoint) 정의

and the corresponding

3) Ideas and Service —— 문제에 대한 답

Concepts

* 컨셉 프로토는 Checkup 발표부터 평가

✓ Midterm Presentation

Mainly about three things

1) Research findings and insights

translated into

2) Design question

} 50%

and the corresponding

3) Ideas and Concepts — 40%

Presentation — 10%

✓ **Midterm Presentation**

1) Research findings and insights

Did you choose the “right” methods? (appropriate selection)

Application of the methods (planning and execution)

Quality of analysis (the depth of insights)

Synthesis of data (within and between methods)

Novel findings (beyond what we know)
혁신의 열쇠

✓ **Midterm Presentation**

2) Design question

The right level of scoping (i.e., granularity)

Grounded by research?

Priority and hierarchy among questions

✓ Midterm Presentation

3) Ideas and Concepts

Rationale behind selecting the concept

~~Multiple concepts~~ ~~(need two or three)~~ At least one concept

Function and task definition

Feasibility of the enabling AI technology 가능하긴 한가?

Cohesiveness with research

Innovativeness, novelty, creativity (USP)

Do you have a chance to solve the problem? 문제가 해결되긴 해?

✓ **Midterm Presentation**

4) Presentation

Clarity

Engagement

Professional manner

Service Concepts

Service Concepts

✓Concepts contain **selected** **theme, notion, or idea** with

선택된 아이디어!

the purpose of **communicating** **a design vision**

about a product or service

(Hartson, Pyla, 2012)

Service Concepts

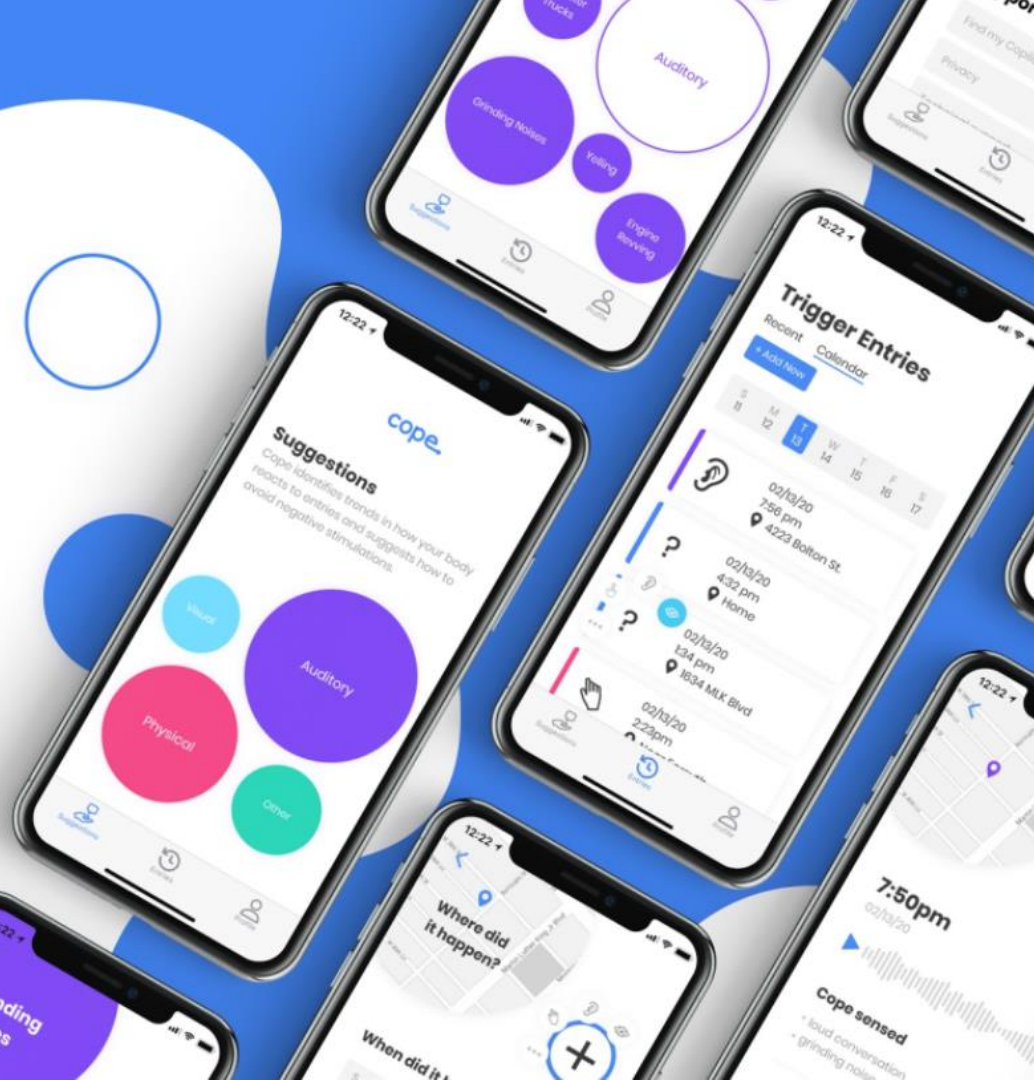
✓ Preliminary concepts are **candidate concepts** to be reviewed and analyzed by clients, peers, executives, developers, engineers, and your target users

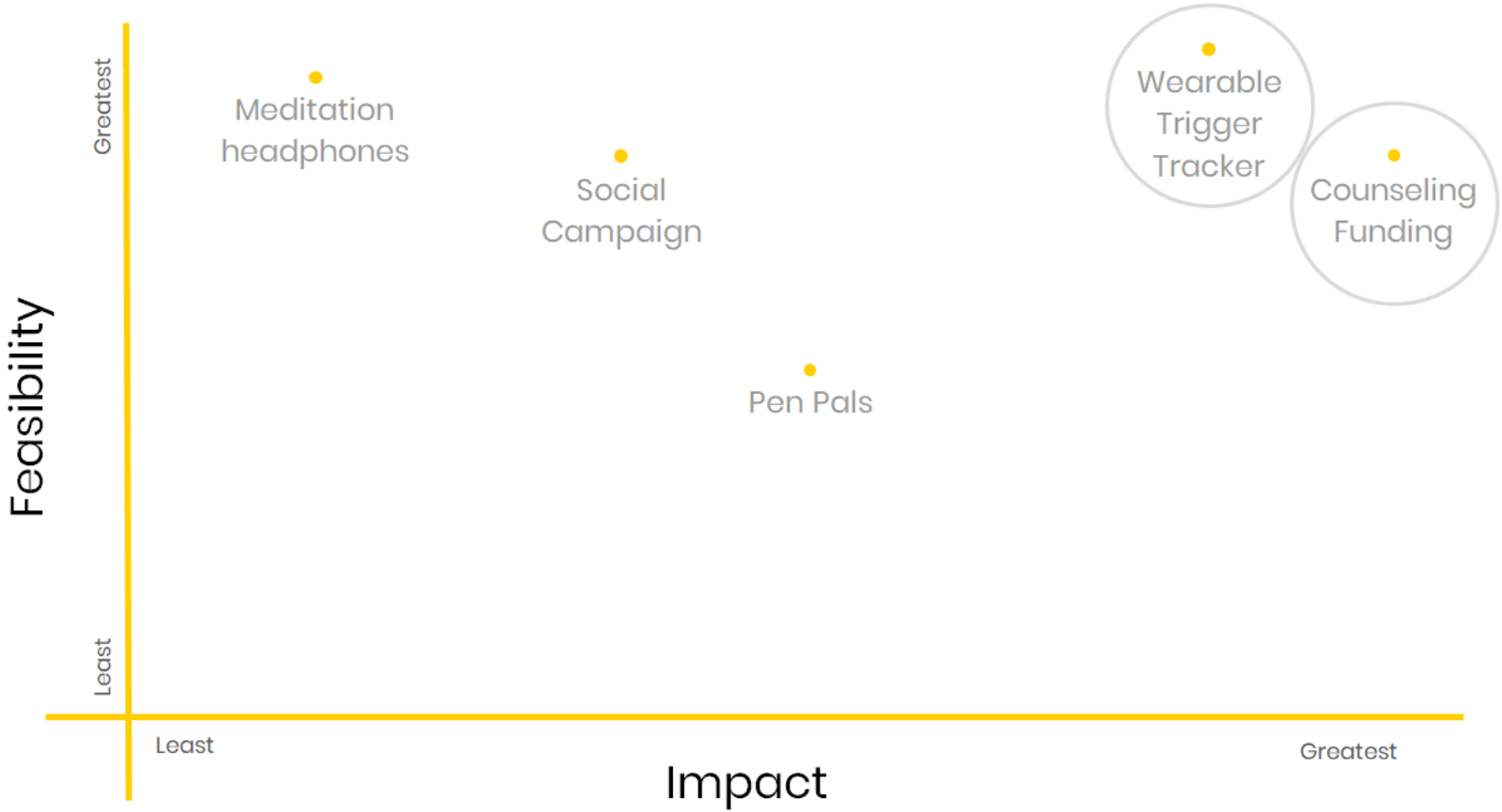
✓ **One will be chosen** to become the final concept

(Hartson, Pyla, 2012)

cope

Let us help you gain control. Let us help you learn.




Feasibility
Greatest
Least
Meditation headphones
Social Campaign
Pen Pals
Wearable Trigger Tracker
Counseling Funding
Least
Greatest
Impact

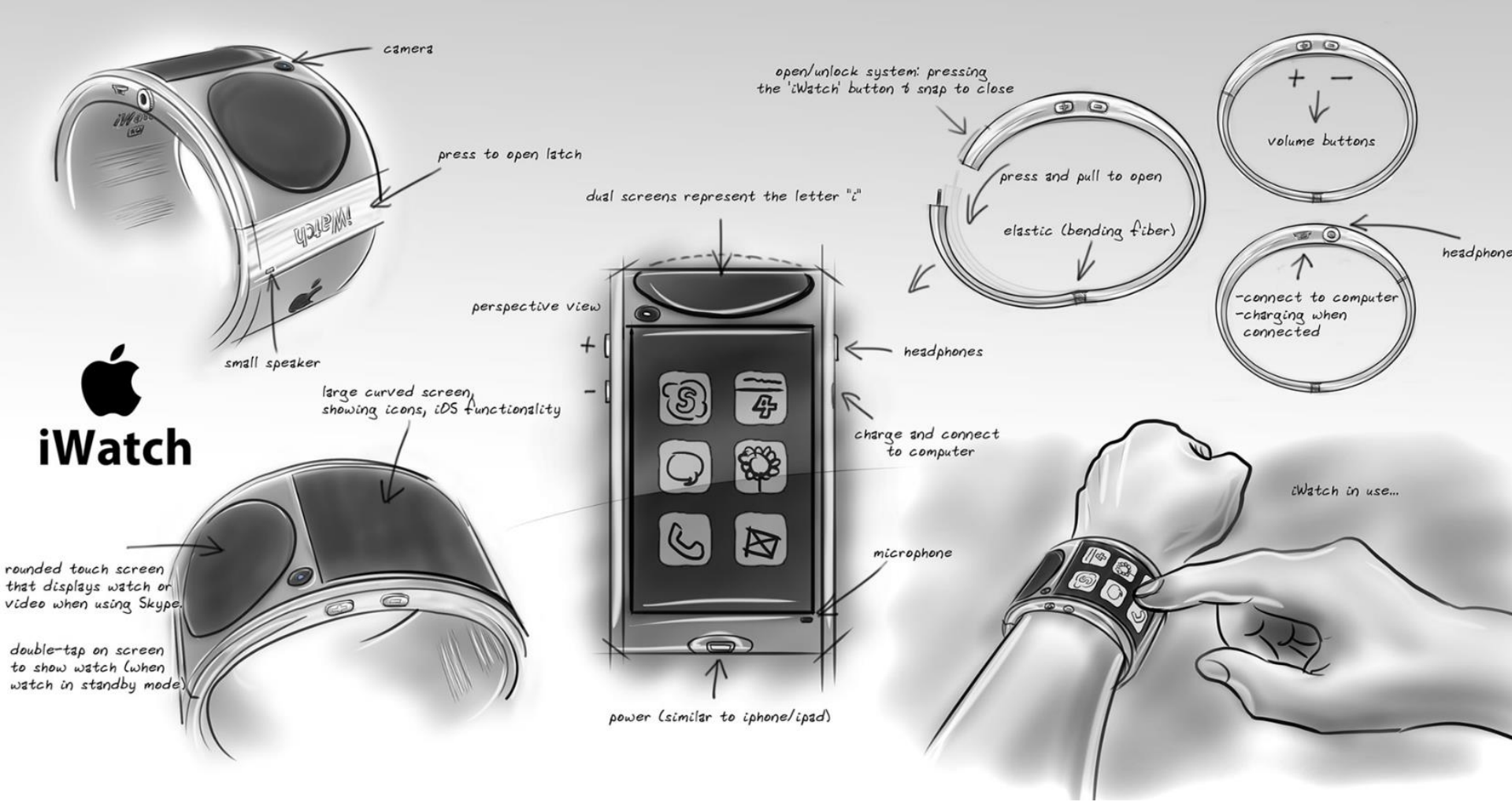
Service Concepts

✓ Your conceptual design, in a form of **sketches**, **storyboard**, and **description**, should convey the group's selected idea

and!

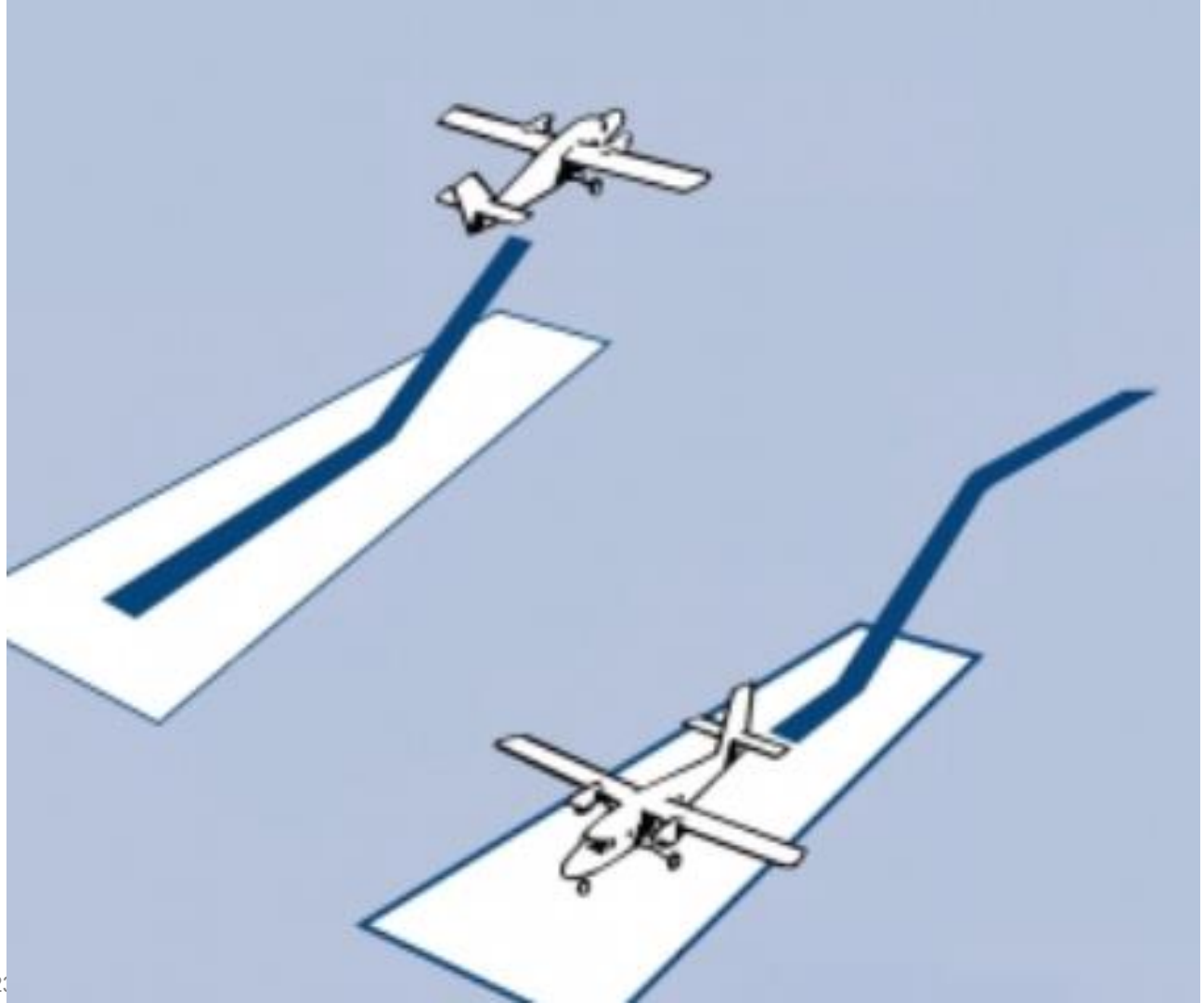
(Hartson, Pyla, 2012)

Sketches



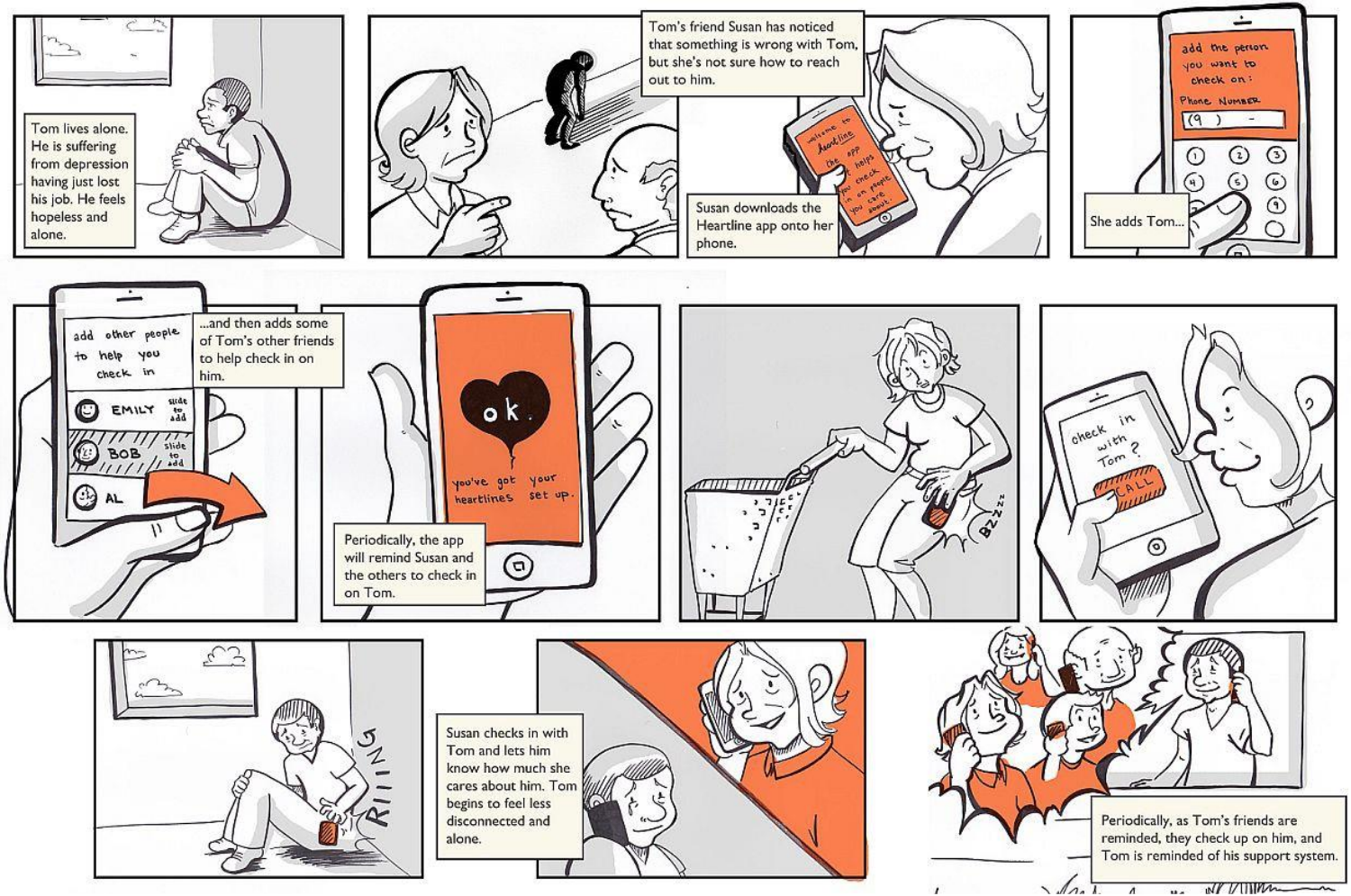
(Reference : iWatch by Jivaldi)

Sketches may include Features, Functions, and User Tasks



Define your
Functions and User Tasks
“very clearly”
in the midterm presentation

Storyboards



(Reference : Chelsea Hostetter, Austin Center for Design)

Why storyboards?

✓ Storyboards create **causation**

because stories involve a form of **chronological narrative**

(i.e., first this happens, next this happens)

✓ People are quick to assign causality

You are always looking for **causation**

Why storyboards?

- ✓ Stories aren’t just for fun
- ✓ Using stories will make it understandable, interesting, and memorable and adds context
- ✓ Use your story and involve the audience in the mid-term presentation!

가성비

Cost-effective

AI Service Research

Out to the field:
Interview and Survey

In-depth Interview

IDI (In-depth Interview)

Strength

Acquiring **Rich**, **Deep**, and **Sincere information**

Also, more **Flexible**

Pressure on subject for **Immediate** response

Face-to-face – can pick up not only verbal

IDI (In-depth Interview)

Weakness

Takes time

Takes money

Not so good to achieve quantitative data from large number of people

IDI (In-depth Interview)

How long?

Typically 30 – 45 minutes (but an hour is possible)

Individual differences on interviewee

Consider the effect of fatigue so you should be observant

IDI (In-depth Interview)

Ideal interview takes place with

Two UX researchers and one user

First person “leads” (ask questions and guide)

Second person takes notes

IDI (In-depth Interview)

How many interviewees?

There is actually a science about this

Number of # is determined by

Heterogeneity of the group

Interview structure and content

Researcher’s skills and experience

Sampling technique

..

But the key is saturation of knowledge

You stop when you sense that you finally see it

See what?


I SEE YOU

You see **the pattern** in the interviewee's experiences

Nevertheless, a good rule of thumb is

between

12 (Guest et al., 2006) to 15 (Bertaux, 1981)

IDI (In-depth Interview)

- **Always the interview script**
- **Always conduct one trial run**
- **Neutral manner (Don’t you think, ….. Right?) (X)**
- **Protect the interviewee (consent)**

Informed Consent Form Example

IDI (In-depth Interview)

Prepare an interview script

Always be mindful of the **higher goals**

Aim for a combination of

Opened ended and **behavioral** (i.e., contextual inquiry)

IDI (In-depth Interview)

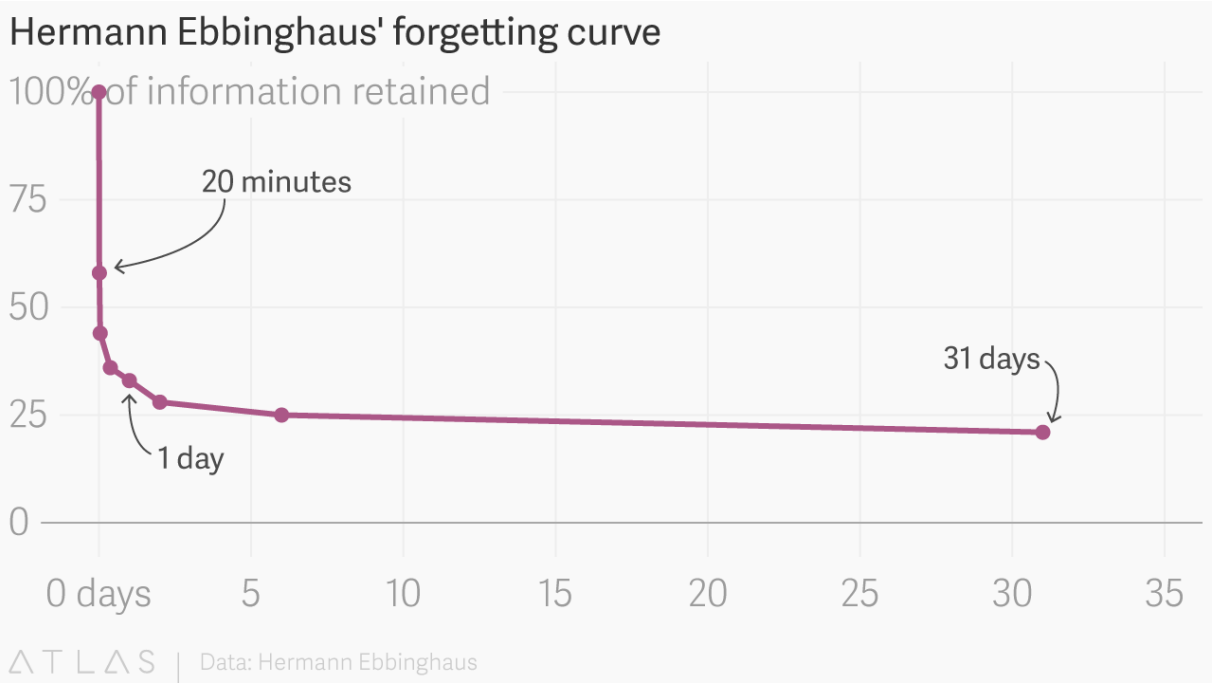
Allow space for follow-up

IDI (In-depth Interview)

Always **download** right away

Or at least before the end of day

IDI (In-depth Interview)



IDI (In-depth Interview)

Affinitize data if necessary

Connect to your **persona design** if necessary

Survey & Questionnaires

Survey

Difference with questionnaires

Survey

Cost-effectiveness

Generalizability

Reliability

Inflexibility

Lack of depth

Survey

Open-ended Questions vs. Close-ended Questions

Qualitative data vs. Quantitative data

Closed Questions

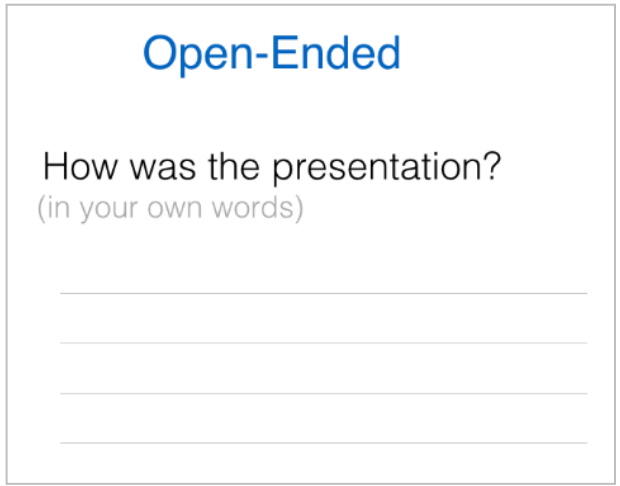
Can get quantitative data
Doesn’t give you the context, the motivation, the cause

29. How well do the teaching styles of your child's teachers match your child's learning style?

- ○ Not well at all
- ○ Mildly well
- ○ Fairly well
- ○ Quite well
- ○ Extremely well

Open Questions

Qualitative data about a user's behaviors and actions
Typically a text box to explain the cause
Takes more time to analyze



Survey

Aim for neutral questions

Leading questions won’t add much value

And can be even intimidating

Survey

One question at a time

Avoid questions that contain two concepts.

It’s confusing and may degrade your data.

Easy and Simple Questions

Keep the language, meaning, and context clear for all the questions.

Make sure your using the language that your target user group can understand

Always run a pilot first

Survey

Easy and simple questions

If you make the user start thinking about your wording . . . or Googling

Survey

You may influence by how you ask

This matters more than you think.

How would you rate product X? vs.

What do you really love about X?

Length of Survey

Focus on time, not on the number of questions

Generally keep it short

Research shows that quality declines on surveys

that are longer than **20 minutes**, so aim for

15 ~ 20 minutes to complete

Sample Size

Question is whether your results are representative

For a **95% confidence level,**

(i.e., only a 5% chance of your sample results

differ from the population average)

Sample Size

Sample Size (*N*)	Margin of Error (fraction)
10	0.316
20	0.224
50	0.141
100	0.100
200	0.071
500	0.045
1000	0.032
2000	0.022
5000	0.014
10000	0.010

5%

Statistically Significant!

5 Point Likert Scale vs 7 Point Likert Scale

- **FIVE – POINT**
- A 5 - point Likert-type scale was used to increase response rate and response quality along with reducing respondents' "frustration level" (Babakus and Mangold 1992). **Source Buttle, F. (1996).**
- A few researchers have, however, reported higher reliabilities for five-point scales (**Jenkins & Taber, 1977; Lissitz & Green, 1975; McKelvie, 1978; Remmers & Ewart, 1941),**
- A five-point scale rather than a seven-point scale was chosen for a number of reasons, one being that it became possible to compare reliability coefficients with other research using five-point Likert Scales. **Saleh, F., & Ryan, C. (1991).**
- **Cox (1980)** concluded that the ideal number of item alternatives seemed to be centered on seven, with some situations calling for as few as five or as many as nine. Also of importance was that an odd number of alternatives, i.e., allowing for a neutral response, were preferable. **Cox III, E. P. (1980).**
- Previous research has found that a five-point scale is readily comprehensible to respondents and enables them to express their views (**Marton-Williams, 1986).**
- The literature suggests that five-point scale appears to be less confusing and to increase response rate (Babakus and Mangold, 1992; Devlin et al., 1993; Hayes, 1992). It has also been suggested that a five-point scale is more appropriate for European surveys (Prentice, 1998). **Source Bouranta, N., Chitiris, L., & Paravantis, J. (2009)**.
- With a Five - point scale, it is quite simple for the interviewer to read out the complete list of scale descriptors ('1 equals strongly disagree, two equals disagree …'). **Dawes, J. G. (2008).** Do data characteristics change according to the number of scale points used? An experiment using 5 point, 7 point and 10 point scales. International journal of market research, 51(1).
- **SEVEN – POINT**
- **Symonds (1924)** was the first to suggest that reliability is optimized with seven response categories, and other early investigations tended to agree (see Ghiselli, 1955, for a comprehensive review of early research). **Source Colman, A. M., Norris, C. E., & Preston, C. C. (1997).**
- **Miller (1956) a**rgued that the human mind has a span of absolute judgment that can distinguish about seven distinct categories, a span of immediate memory for about seven items, and a span of attention that can encompass about six objects at a time, which suggested that any increase in number of response categories beyond six or seven might be futile. **Colman, A. M., Norris, C. E., & Preston, C. C. (1997).**
- Lewis (1993) found that 7-point scales resulted in stronger correlations with t-test results. **Lewis, J. R. (1993).**
- Seven-point Likert scales appear to be more suited to electronic distribution of usability inventories. **Finstad, K. (2010).**
- Research confirms that data from Likert items (and those with similar rating scales) becomes significantly less accurate when the number of scale points drops below five or above seven. **Johns, R. (2010).** Likert items and scales. Survey Question Bank: Methods Fact Sheet, 1.
- In the light of findings, there is some support for seven-point scales, but the popularity of five-point scales seems to be less justified. **Preston, C. C., & Colman, A. M. (2000).**

Likert Scale Labels

Group discussion