

Grading steps - design review

- Tests
 - Test color gradation, is reusing green and red a good idea?
 - 1. Are the icons clear enough?
 - Preparation
 - Execution
 - Results
 - Expert test
 - Comments
 - Package state
 - Comments
 - Predefined remarks
 - Comments
 - Results form
 - 2. Test damage screen. All in one or separate?
 - Users/learning effect
 - Preparation
 - Execution
 - Results
 - Combined questions
 - Comments
 - Separate questions
 - Comments
 - Which option has preference?
 - Results form
 - 3. Are the checkbox questions clear?
 - Users/learning effect
 - Preparation
 - Execution
 - Results
 - Were the right options selected?
 - Did user select "Next"?
 - Results form
 - 4. Is it clear that the colors indicate the scale from good to bad?
 - Users/learning effect
 - Preparation
 - Execution
 - Results
 - Did user pick the right color?
 - Results form
 - 5. Rate how much better do you think this design is compared to the old one
 - Users/learning effect
 - Preparation
 - Execution
 - Results
 - Comments
 - Results form
 - 6. Is it clear how to get more information about a product?
 - Users/learning effect

- Preparation
- Execution
- Results
 - Comments
- Results form
- 7. How would you pause this process?
 - Users/learning effect
 - Preparation
 - Execution
 - Results
 - Comments
 - Results form
- 8 Is the selected state that is shown going back clear enough
 - Users/learning effect
 - Preparation
 - Execution
 - Results
 - Results form
- Prototype
- General preparation & execution
- Planning
- Preparation Wijchen
- Planning
- Follow up

Tests

~~Test color gradation, is reusing green and red a good idea?~~

Decided that we do not need to test this

1. Are the icons clear enough?

Preparation

1. Print the icons for the 'Expert test' and the labels, labels and icons on separate cards
2. Same for 'Package state' and 'Predefined remarks'

Execution

1. Ask at least 2 users to match Expert test icons and labels
2. Same for 'Package state' and 'Predefined remarks'. They may be the same test users as in 1.

If the users have no problems matching icons and labels, then they both convey the same message and then it is safe to say that the icons strengthen the labels.

Results

4 graders

Expert test

- 4 out of 6. 25%
- 6 out of 6. 75%

Comments [↗](#)

- Heeft redelijk wat tijd nodig om tot een antwoord te komen. Niet functioneel heel duidelijk. Lastiger dan gedacht
- Twijfelt over beschadigd en ernstig beschadigd. Binnen minuut klaar
- Mixed ernstig en beschadigd
- Neemt de tijd, twijfelt onderweg. Afgeleid door de drie streepjes. Lijkt op barst. Ernst van schade is lastig

Package state [↗](#)

- 4 out of 6. 25%
- 6 out of 6. 75%

Comments [↗](#)

- Duidelijke verschillen. Neemt rustig de tijd
- Snel klaar
- Heel snel klaar maar ongeopend en geopend verwisselt. Gewend om naar plaats te kijken
- Heeft moeite met begrijpen van teksten. Lijkt teveel op elkaar

Predefined remarks [↗](#)

- 6 out of 6. 100%

Comments [↗](#)

- Mist kruisje door cups
- Makkelijk

Results form [↗](#)



2. Test damage screen. All in one or separate? [↗](#)

Do graders prefer to use one screen with all options or separate questions per option. Or is there no preference

Users/learning effect [↗](#)

The users may have done test 1.

Preparation [↗](#)

1. An interactive prototype with 4 separate damage steps (front/top, other, display, free text) for an iPhone with a small scratch on the display only.
2. Same but now for all damage integrated in 1 step

Execution [↗](#)

1. We ask a user to fill in the damage for prototype 1.
2. We ask the same users to fill in the damage for prototype 2
3. We ask them what they prefer
4. Repeat with more users, but alternate the order of prototype 1 & 2

Results [↗](#)

Combined questions [↗](#)

Did user complete the task with combined steps?

- No. 100%
- Yes. 0%

Comments [↗](#)

- Kiest top ipv scherm in eerste vraag. Niet duidelijk dat accessoires overgeslagen kan worden (heeft geen "geen"). Vind verder stap uiteindelijk wel
- Kiest voor kras op product. Niet duidelijk dat Next voor "geen keuze" is. Weet niet wat er nog komt
- Kiest voor kras op product. Accessoires vraag niet duidelijk. Kiest toch voor schade (1) ipv next
- Kiest in eerste vraag de schade. Klik na uitleg op geen. Bij acc. wordt Next gekozen

Separate questions [↗](#)

Did user complete the task with separate questions?

- Yes. 50%
- No. 50%


Comments [↗](#)

- Kiest in eerste instantie voor verkeerde optie

Which option has preference? [↗](#)

- Separate. 0%
- Combined. 100%

Results form [↗](#)

 Page not found

3. Are the checkbox questions clear? [↗](#)

Is it clear that multiple options can be chosen? Is it clear that you have to click next for checkbox questions?

Users/learning effect [↗](#)

These users may not have done test 2.

Preparation [↗](#)

1. Bring a product
 - a. Jabra wireless earbuds in original box (I have them)
2. Make a prototype with a new (non scaled) radio button question the graders don't know yet: "What is the brand of the product?". (this servers to create the bias that only 1 option is possible)
3. Append a new checkmark question: "Select what is applicable"
 - a. The product is black
 - b. The product has an Azerty keyboard
 - c. The original box is present
4. Append another dummy question, contents not important.

Execution [↗](#)

1. Give the user the Jabra wireless earbuds in the box
2. Let the user answer question 1.
3. Let him answer question 2 (what is applicable?).
4. He will stay on the same page. Observe if he checks 1 or 2 answers.

5. Observe if he clicks on 'Next' by himself.

Results [↗](#)

6 graders

Were the right options selected? [↗](#)

- Yes. 66%
- No. 33%

Did user select "Next"? [↗](#)

- Yes. 100%

Results form [↗](#)

 [Page not found](#)

4. Is it clear that the colors indicate the scale from good to bad? [↗](#)

Users/learning effect [↗](#)

They may have done test 3 (has no scales), but not test 2 (has scales).

Preparation [↗](#)

1. "Package state" screen with blurred text and images, range from green to red

Execution [↗](#)

1. For screen 1 we ask at least 2 users what they would click on if the package was in a really bad condition and would have to be replaced.

Hypothesis "a scale from green to red is a good idea" is accepted if users consistently pick options on the right side of the middle.

Results [↗](#)

6 graders

Did user pick the right color? [↗](#)

- Yes (100%)

Results form [↗](#)

 [Page not found](#)

5. Rate how much better do you think this design is compared to the old one [↗](#)

Users/learning effect [↗](#)

Users may have experience, but that might influence their judgement. Would be very nice if we can ask additional users without any testing experience.

Preparation [↗](#)

1. Make a screenshot of the "old package state" screen
2. Add the new design of "package state" to the prototype

Execution

1. Show a user the old design
2. Show a user the new design
3. Ask for a comparison
 - new is much worse than old
 - new is worse than old
 - new is equal to old
 - new is better than old
 - new is much better than old

Results

14 graders

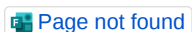
Is the new design better or worse?

- New is worse than the old one (7%)
- New is equal to old (43%)
- New is better than old (29%)
- New is much better than old (21%)

Comments

- Lastig kiezen door gewenning
- Leest toch tekst wel
- Kent scherm al. Werkt met muis
- Kan sneller zien wat ze moet doen. Vrolijke kleuren
- Gewend aan de oude
- Alleen tekst. Nieuw is duidelijker
- Nieuwe is makkelijker om te leren
- Nieuw is overzichtelijker. Moet teveel lezen in oude. Tekst in knoppen duidelijker
- Snelheid maakt niet uit. Wordt misschien te makkelijk. Automatische piloot
- Nieuw is makkelijker. Kiest vaak hetzelfde in deze vraag. Leest teksten niet meer. Nieuw nodigt meer uit
- Veel duidelijker, overzichtelijker. Kleur is duidelijk. Mooi
- Verandering is lastig. Nieuw is prettiger, vriendelijker
- Nieuw is beter. Duidelijker, ook door kleuren
- Gaat automatisch. Geen voorkeur

Results form



6. Is is clear how to get more information about a product?

Users/learning effect

They may have experience with other tests. (but it may be wise to remove the product area from the prototype everywhere except for the screens of this test so that they did not notice it before!)

Preparation [↗](#)

1. Prepare a screen with a clickable product area (right top)

Execution [↗](#)

1. Show the screen with a clickable product area
2. Ask the user what he would do if he wants to know the previous grade of this product
3. Open LP details if they don't know how
4. Ask what they would do to continue grading

Results [↗](#)

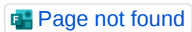
Did user find the option to get more information?

- No. 50%
- Yes. After a while. 25%
- Yes. 25%

Comments [↗](#)

- Probeert op EAN te klikken
- Kiest Next button. Had Info wel gezien
- Opent meestal nieuwe tab voor deze taak

Results form [↗](#)



7. How would you pause this process? [↗](#)

Users/learning effect [↗](#)

We need users that did not do test 6.

Preparation [↗](#)

1. Use the prototype of 6.

Execution [↗](#)

1. Show a screen with a clickable product area (right top)
2. Ask the user how he would pause this grading step
3. If stuck: show the LP details. Ask again how he would pause.

Results [↗](#)

7 graders

Was the user able to pauze the grading process?


- No (28%)
- Yes. After a while (28%)
- Yes. After being pointed to license plate view (28%)
- Yes (14%)

Comments [↗](#)

- Zou scherm laten staan

- Moet goed zoeken

Results form [↗](#)

 [Page not found](#)

8 Is the selected state that is shown going back clear enough [↗](#)

Users/learning effect [↗](#)

They may have experience with other tests.

Preparation [↗](#)

1. Create an interactive prototype with a radiobutton question 'What is the brand' (Philips/Jabra/Sony)
2. Add an additional dummy question with an interactive back button
3. If you click 'Back', you should see 'What is the brand' with 'Philips' selected

Execution [↗](#)

1. Give the user the Jabra headset
2. Ask the user to answer to answer the grading question
3. Ask the user to go back
4. Ask him if they notice anything here ('Jabra' is not selected, but 'Philips')


If the user quickly sees that something is wrong, that means that he does notices the selected answer.

Results [↗](#)

10 graders

- Nothing special noticed. 10%
- Jabra was selected instead of Phillips (correct). 90%

Results form [↗](#)

 [Page not found](#)

Prototype [↗](#)

 [Touch friendly grading - prototype](#)

General preparation & execution [↗](#)

- ✓ [Prototype](#) bouwen (DJ)
- ✓ Prepare forms to note the answers on all the tests (Dennis)
- ✓ Grading station reserveren voor halve dag (Dennis)
- ✓ Testers reserveren met variërende ervaring (Dennis)
- ✓ Bring Jabra headset (DJ)
- ✓ Test de test (Dennis + DJ) in Amsterdam (test hardware setup en test zelf)

- ✓ Hardware test Wijchen
- ✓ 1 doet test & instructies (DJ)
- ✓ ander noteert en haalt mensen op (Dennis)
- ✓ Dennis maakt foto's voor een presentatie
- ✓ Prepare 'spiekbriefje'

Planning [↗](#)

May 24th in Wijchen. Start at 10:00

Preparation Wijchen [↗](#)

- ✓ Provide a list of graders that will participate. Probably one from each team in every group
- ✓ Check if teamleads want to do the test as well

Planning [↗](#)

Group	Schedule	
Group A (3)	1 → 2 → 6 (→ 8) → 5	
Group B (3)	3 → 7 → 8 (→ 4) → 5	test 4 only when time available
Group C	5	

Estimation 15-30min per user, half day will probably suffice

Ideally we have some variation in experience in the testers

Follow up [↗](#)

- ☐ Share test results with Jeannette and potentially with teamleads
- ☐ Determine next steps