Good bye conversion rate
a smarter way to optimising Search Engine Results Pages

////// IS-SWIS 2017, Rotterdam, February 9, 2017////

MartinGaedke.com
VSR.Informatik.TU-Chemnitz.de
Search Engine Result Pages (SERPs)
(simplified) Motivation

<table>
<thead>
<tr>
<th>Evaluation Method</th>
<th>Efficient?</th>
<th>Effective?</th>
</tr>
</thead>
<tbody>
<tr>
<td>User Testing</td>
<td>☒</td>
<td>☑</td>
</tr>
<tr>
<td>Usability Inspection</td>
<td>☒</td>
<td>☑</td>
</tr>
<tr>
<td>Split Testing</td>
<td>☑</td>
<td>☒</td>
</tr>
</tbody>
</table>
Split Testing

Version A

Buy this

Version B

Buy this
The problem with Split Testing

*Conversion Rate:* proportion of visitors to a website who take action due to subtle requests from marketers, SEOs, (semantic) data providers etc.

Conversion rate = \[
\frac{\text{Number of Goal Achievements}}{\text{Visitors}}
\]

(source: Wikipedia.com)
Version A: Getting the User’s Zip Code
$$\neq$$ Usability

Jakob Nielsen: Putting A/B Testing in Its Place,
http://www.nngroup.com/articles/putting-ab-testing-in-its-place/
Goal: to leverage the advantages of split testing AND to facilitate effective usability evaluation
Idea:
Rethink the target measure from the users point of view (i.e. development & user – not sales):
Usability as a measure for split testing
Idea: Infer quantitative measure of usability from user interactions.
Usability-based Split Testing

Interface A

Questionnaire

Model for Usability

WaPPU
Usability System
(Naive Bayes classifier for the learning)

Interface B

[WaPPU – Usability based A/B-Testing
Speicher, Both, Gaedke, ICWE2014]
Usability-based Split Testing

Interface A → TRAINED Model for Usability → Interface B

WaPPU Usability System

[ WaPPU – Usability based A/B-Testing Speicher, Both, Gaedke, ICWE2014 ]
**Usability Model allows for:**

\[ f(\text{cause}) = \text{score} \]

*(cause is part of usability)*

### Interface “A” vs. Interface “B”

<table>
<thead>
<tr>
<th></th>
<th>Interface “A”</th>
<th>Interface “B”</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Overall Usability</strong></td>
<td>81.3% ± 37.2%</td>
<td>18.8% ± 37.2%</td>
</tr>
<tr>
<td>informativeness</td>
<td>0.6 ± 0.5</td>
<td>-0.6 ± 0.5</td>
</tr>
<tr>
<td>understandability</td>
<td>0.6 ± 0.5</td>
<td>-0.6 ± 0.5</td>
</tr>
<tr>
<td>confusion</td>
<td>0.6 ± 0.5</td>
<td>-0.6 ± 0.5</td>
</tr>
<tr>
<td>distraction</td>
<td>0.6 ± 0.5</td>
<td>-0.6 ± 0.5</td>
</tr>
<tr>
<td>readability</td>
<td>0.6 ± 0.5</td>
<td>-0.6 ± 0.5</td>
</tr>
<tr>
<td>infDensity</td>
<td>0.6 ± 0.5</td>
<td>-0.6 ± 0.5</td>
</tr>
<tr>
<td>accessibility</td>
<td>0.6 ± 0.5</td>
<td>-0.6 ± 0.5</td>
</tr>
</tbody>
</table>

*Interface “B” is significantly worse than Interface “A”.*
Limitations?
Sure!

\[ f(cause) = score \]
this function is \textit{not bijective} 

(Not possible to infer \textit{optimizations} from usability scores)
How to overcome the limitations?

\[ f(cause) = score \]
Catalog of best practices
a.k.a. Set of causes and optimizations
SERP Optimization Suite (S.O.S.)

WaPPU for evaluation

Catalog of best practices

\( f_u \) (score) = \{C, C'\}

\( u \) = usability item number

\( C \) = potential causes

\( C' \) = countermeasures
S.O.S. example: "We have a problem with Distraction"
A. Informativeness
   I. bad index quality (desired result[s] not present on page)
   II. bad ranking quality (desired result[s] not present or ranked too low)
   III. desired result not clearly identifiable:
      a. inappropriate title and/or abstract
      b. too many other results
      c. too much content other than results
   2. ⑨ provide search suggestions/related search terms
   6. clarify layout:
      b. ①④⑧ clearly separate results from other content such as ads (or remove the latter) [2–1]
      c. ①②③ reduce amount of content other than results
      e. ④ clearly separate title from abstract [3–0]

B. Understandability
C. Confusion
   II. too many irrelevant results [3–0]
   III. too much content other than results
   IV. too many advertisements which are poorly marked as such; no clear separation between advertisements/spONSponsed results and real results [6–0]
   2. ⑧ clearly highlight results and mark other content as such
   5. ①②③ reduce amount of content other than results

D. Distraction
   I. too much content other than results
   III. too many images
   IV. non-results more salient than results [2–1]
   V. overloaded results (e.g., displaying secondary information, social media buttons etc.)
   1. ①②③ reduce amount of content other than results
   3. ⑤ reduce amount of images
   4. ③④⑤⑥⑦⑧ ensure that results are more salient than other content [3–0]
   5. clarify presentation of results:
      a. ② reduce to: title, URL, abstract

E. Readability
   I. wrong font size or character spacing (too small/too large) [9–0]
   II. wrong line height (too small/too large) [6–0]
   V. text not properly grouped (e.g., via white space or separation lines) [3–0]
   VII. inconsistent alignment of results and/or other elements of the page
   1. ⑤ adjust font size or character spacing (or offer according option to reader) [11–1]
   2. ⑤ adjust line height [9–0]
   5. add white space:
      a. ④⑦ between title, URL and abstract of result
      b. ④ between results
      c. ④⑦ between results and other content
   7. ⑥⑦ align results and other elements of the page consistently

F. Information Density
   III. too much content other than results; content that is not related to results [3–0]
   IV. too little white space
   V. missing visual hierarchy with salient results [2–1]
   VI. overloaded results (e.g., displaying secondary information, social media buttons etc.)
   3. ①②③ reduce amount of content other than results
   4. add white space:
      a. ④⑦ between title, URL and abstract of result
      b. ④ between results
      c. ④⑦ between results and other content
   5. ⑥⑦⑧ introduce contrast and visual hierarchy to separate results from content other than results [3–0]
   10. clarify presentation of results:
      a. ② reduce to: title, URL, abstract
   11. ③ remove unnecessary icons/abbreviations or add explaining tooltips [3–0]

G. Accessibility
   I. too much scrolling effort for user:
      a. too much content other than results, especially above results
      d. bad ranking quality (desired result[s] not present or ranked too low)
   II. desired result(s) not immediately identifiable:
      b. missing visual hierarchy [2–1]
      c. missing contrast between results and other content [3–0]
   1. reduce scrolling effort:
      a. ① reduce amount of content other than results, especially above results
   2. better highlight results/improve result presentation:
      b. ⑧ introduce contrast and visual hierarchy to separate results from content other than results [2–1]
      c. ③④⑤⑥⑦⑧ ensure that results stand out against other content
A. Informativeness
   I. bad index quality (desired result[s] not present on page)
   II. bad ranking quality (desired result[s] not present or ranked too low)
   III. desired result not clearly identifiable:
         a. inappropriate title and/or abstract
         b. too many other results
         c. too much content other than results
   2. adjust line height [9-0]
   5. add white space:
      a. ④ ⑦ between title, URL and abstract of result
      b. ④ between results
      c. ④ ⑦ between results and other content
   7. ⑥ ⑦ align results and other elements of the page consistently

D. Distraction
   I. too much content other than results
   III. too many images
   IV. non-results more salient than results [2-1]
   V. overloaded results (e.g., displaying secondary information, social media buttons etc.)
   1. ①②③ reduce amount of content other than results
   3. ② reduce amount of images
   4. ③④⑤⑥⑦⑧ ensure that results are more salient than other content [3-0]
   5. clarify presentation of results:
      a. ② reduce to: title, URL, abstract

E. Readability
   I. wrong font size or character spacing (too small / too large) [9-0]
   II. wrong line height (too small / too large) [6-0]
   V. text not properly grouped (e.g., via white space or separation lines) [3-0]
   VII. inconsistent alignment of results and/or other elements of the page
   1. ⑤ adjust font size or character spacing (or offer according option to reader) [11-1]

F. Information Density
User Study

- 81 participants (~62% male)
- avg. age = 31.08 years
- Task: Solve a problem
“Find a birthday present for a good friend that does not cost more than 50 Euros.”
New York City

NEU! Hotel Mallorca Khao Lak -70% +++ Ramada - Khao Lak - Spanien
Hotel The Westin Leipzig, Leipzig, Sachsen 168 Hotelbewertungen für The Westin Leipzig Angebote
The Westin Leipzig, Urlaub The Westin Leipzig

Verwandte Suchanfragen:
- Mallorca Viertel 14 Tage
- Viertel Auf Mallorca
- Viertelwagen Mallorca
- Viertelwohnung Mallorca
- Viertelorte auf Mallorca
- Hotel Mallorca

Hotel Mallorca Khao Lak - Ihr luxuriöses 4.5 Sterne Hotel
Hotel The Westin Leipzig, Leipzig, Sachsen 168 Hotelbewertungen für The Westin Leipzig Angebote
The Westin Leipzig, Urlaub The Westin Leipzig
http://www.weg.de/hotels#e-westin-leipzig-4209

Martin Gaedke - http://vsr.informatik.tu-chemnitz.de
## Results

<table>
<thead>
<tr>
<th></th>
<th>old SERP</th>
<th>new SERP</th>
</tr>
</thead>
<tbody>
<tr>
<td>informativeness</td>
<td>-0.17 \downarrow &lt;</td>
<td>-0.02 \downarrow &lt;</td>
</tr>
<tr>
<td>understandability</td>
<td>0.34 \downarrow &lt;</td>
<td>0.45 \downarrow &lt;</td>
</tr>
<tr>
<td>confusion</td>
<td>0.30 \downarrow &lt;</td>
<td>0.38 \downarrow &lt;</td>
</tr>
<tr>
<td>distraction*</td>
<td>0.36 \downarrow &lt;</td>
<td>0.62 \uparrow &lt;</td>
</tr>
<tr>
<td>readability</td>
<td>0.45 \downarrow &lt;</td>
<td>0.52 \downarrow &lt;</td>
</tr>
<tr>
<td>information density*</td>
<td>0.04 \downarrow &lt;</td>
<td>0.43 \uparrow &lt;</td>
</tr>
<tr>
<td>accessibility</td>
<td>0.06 \downarrow &lt;</td>
<td>0.07 \downarrow &lt;</td>
</tr>
<tr>
<td>overall usability*</td>
<td>59.86% \downarrow &lt;</td>
<td>67.50% \uparrow &lt;</td>
</tr>
</tbody>
</table>

Martin Gaedke - http://vsr.informatik.tu-chemnitz.de
Concluding thoughts
Conclusions

- S.O.S. significantly improved SERP usability
- Approach developed using HCD/DT
  - obtaining usability scores through WaPPU and
  - recommendations for optimization from a catalog for suboptimal scores
- [http://vsr.informatik.tu-chemnitz.de/demo/SOS](http://vsr.informatik.tu-chemnitz.de/demo/SOS)
Future Work

- Catalogs for other categories of Web pages
  - Creating trained model for web page and site categories
  - Creating more generic and reusable catalogs
  - Reducing the learning time
- Applying the approach to the quality of linked enterprise data
  - Improving: incorrect, incomplete etc. data
  - Dealing with: Co-evolution, Coherence etc.
  - Data Quality → Fit for business use
  - LEDS project: http://www.leds-projekt.de/
Data Quality can be interpreted as the degree to which data fits to current requirements
First Inspiring Results
Output of Fitness results of our Quality Assessment tool with the means of the data quality vocabulary (dqv)

Feel free to visit: www.leds-projekt.de
Thank You!

martin@gaedke.com

VSR.Informatik.TU-Chemnitz.de

@gaedke /gaedke