

The Effect of **Severity Ratings** on **Software Developers' Priority** of Usability Inspection Results

I-USED 2008
Pisa, Italy

Asbjørn Følstad
SINTEF ICT

Content

- Background – a small inconsistency
- A small study – aiming to resolve the inconsistency
- A controversial conclusion - withdrawn

Background I: The importance of impact



High severity
issues?
→



→
Low severity
issues?



Recycle Bin

Background II: Impact defined

$$\text{Impact} = \frac{\text{Redesign!}}{\text{Redesign!} + \text{Recycle Bin}} \times 100$$

Background III:

A small inconsistency

- Sawyer, Flanders, Wixon: Making a difference – the impact of inspections (CHI'96)
- Ten usability inspection cases at Digital Equipment Corp.
- No difference in impact between low, medium or high severity issues (72, 71, 72 %)
- Law: Evaluating the downstream utility of user tests and examining the developer effect (IJHCI, 2006)
- One case of user tests - developing a multilingual digital academic library (v0.85 -> v1)
- Low severity issues given lower priority than medium and high severity issues (26, 42, 47 %)

Background III: **A small inconsistency**

- Sawyer, Flanders, Wixon:
Making a difference – the
impact of inspections (CHI'96)

- Law: Evaluating the
downstream utility of user tests
and examining the developer
effect (IJHCI, 2006)

Relevant study
(UIM), but
surprising
results



Almost relevant
study (not UIM),
but comforting
results

A small study: Details

- Usability inspections
 - Group-based expert walkthrough
 - 4-5 usability experts in each group
 - One group for each case
- Three cases of mobile application development
 - Medical personnel in hospitals
 - Politicians and political advisors
 - Parking wardens
- Evaluation reports
 - 167 usability issues across all cases
 - Rated as cosmetic, serious, critical
- Impact measures
 - Collected three months following report handover
 - Change: (High priority)
 - No change (Medium or low priority – or wrong)



A small study:

Results

CASE 1	Not class.	Cosmetic	Serious	Critical
No change	6	3	6	9
With change	1	2	8	6
Impact ratio	14 %	40 %	57 %	40 %

CASE 2	Not class.	Cosmetic	Serious	Critical
No change	9	15	7	7
With change	4	5	9	5
Impact ratio	31 %	25 %	56 %	42 %

CASE 3	Not class.	Cosmetic	Serious	Critical
No change	31	13	13	0
With change	1	2	1	1
Impact ratio	3 %	13 %	7 %	100 %



A small study: Results

ALL CASES	Not class.	Cosmetic	Serious	Critical
No change	46	31	26	16
With change	6	9	18	12
Impact ratio	12 %	23 %	41 %	43 %



A small study:

Conclusion

- Low severity usability issues with significantly lower impact than medium and high severity issues
- Sawyer et al. contradicted; Law supported
- Why diverging results between the studies?
 - Herzum (2007): Development stage may affect impact
 - Early phase: Higher impact -> Less difference between low and high severity issues?
 - Later phase: Lower impact -> Larger differences?
- Both Law's study and the small study target later phase
- Phases targeted in Sawyer et al. not known.

A controversial conclusion: **Withdrawn**

Discussion point: Was I right
In withdrawing this conclusion?

