

Problems of Consolidating Usability Problems

Effie Lai-Chong Law

**ETH Zürich, Switzerland/
University of Leicester, UK**



Ebba Thora Hvannberg

University of Iceland

ETH

Eidgenössische Technische Hochschule Zürich
Swiss Federal Institute of Technology Zurich

Motivation

- Multiple users/analysts identify different sets of usability problems (UPs)
 - Consolidating these sets into a master list
 - Provide neat information for development team, facilitating redesign
 - Enhance validity of the UPs identified
 - UP consolidation: common practice, little is known about:
 - **HOW** it is exactly done
 - **WHAT** impact it can have on final evaluation outcomes and eventually on system redesign
 - Problem extraction on **individual** basis rather than problem merging on a **collaborative** basis (cf. Nielsen's UE, Hertzum's evaluator effect)
-

UP Consolidating Process

- First phase – **Filtering**, to eliminate any duplicate **WITHIN** a list of UPs
 - Second phase – **Merging**, to combine UPs **BETWEEN** different lists
 - To retain unique, valid UPs
 - To discard unique, invalid UPs
 - The actual practice of UP consolidation is largely **open, unstructured and unchecked.**
-

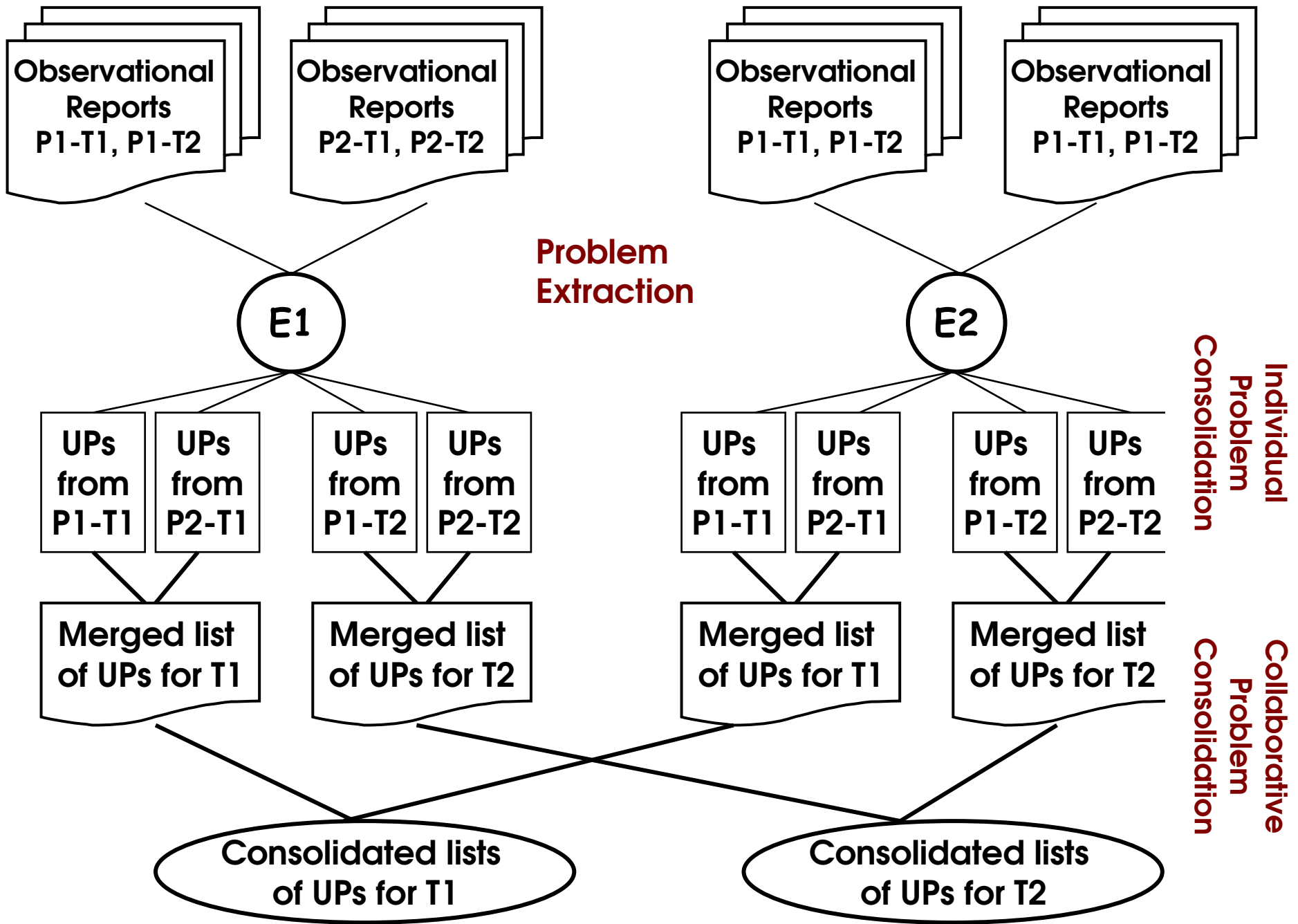
Goals

- **To examine the impact of the UP consolidation process (quantitative approach)**
 - **To understand the mechanism underlying the consensus building process (qualitative approach)**
-

Method

Participants

- 5 pairs of university students (UK), majored in computer science, 3rd year Bachelor or Master level
 - Reasonable knowledge in HCI and some practical experience in usability evaluation
-



Tendency to Merge

Individual Filtering

	Merged	Discarded	Retained
T1	39%	13%	48%
T2	51%	10%	39%

- Problem types vs. Problem instances

- Avoiding direct confrontation

Collaborative Filtering

	Merged	Discarded	Retained
T1	81%	10%	9%
T2	77%	15%	8%

- Receptive to suggestion, no actual economic loss or personal gain

Change in Severity

	Severity		Confidence	
	T1	T2	T1	T2
DEC	4 (10%)	3 (7%)	6 (15%)	4 (10%)
SAME	20 (53%)	29 (71%)	15 (40%)	18 (44%)
INC	14 (37%)	9 (22%)	17 (45%)	19 (46%)

Individual

	Severity		Confidence	
	T1	T2	T1	T2
DEC	2 (5%)	2 (7%)	2 (5%)	3 (11%)
SAME	23 (52%)	16 (57%)	22 (50%)	13 (46%)
INC	22 (43%)	10 (36%)	19 (45%)	12 (43%)

Collaborative

Discussion

- **The lenience towards merging UPs in the collaborative setting**
 - the high merging rate
 - combined UPs that had not been merged in individual sessions to merge with the partner's
 - **Social pressure to reach consensus**
 - **As a result of the merging process in the collaborative setting:**
 - severity ratings of UPs tend to inflate
 - the number of UPs tends to deflate excessively
 - **Confidence levels, in which personal experience plays a role, do not fluctuate with the merging process**
-

Impact

- Invalid severity ratings presumably lead to the fixing of less urgent UPs.
 - the quality of the system may still be undermined by more severe as well as more urgent UPs.
-

Future Work

- to look into relevant theories on similarity, communication, argumentation, and social interaction
 - to extend our empirical studies by systematically comparing merging through negotiation versus merging through authority
 - to develop objective measures of the cost-effectiveness of such methods
 - to contribute to research work on downstream utility
-

Thank you!
Questions?
