Investigation of Indirect Oral Operation Method for Think Aloud Usability Testing

Masahiro Hori, Yasunori Kihara, Takashi Kato

Kansai University

Outline

- Backgrounds
  - Interactive prototype
  - Concurrent think-aloud protocols
- Approach to continuous verbalization
  - Oral operation method
- Comparative evaluation
  - Easiness of operation with the oral operation
  - Number of utterances collected
- Conclusions

Background: Prototyping

- Used for the evaluation of design ideas
  - usability & user experiences
- Interactive prototypes
  - helpful for testing enhanced input capabilities (touch screen & sensors)
  - may not achieve operational performance expected in the final product
    - due to slow or inaccurate response of the software

Example of Interactive Prototype: Touch-Screen Digital Camera

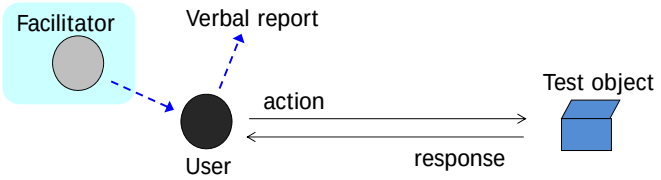
- Running on a tablet PC

- Reported flaws may come from
  - not only intrinsic problems of the artifact,
  - but also insufficient operational performance of the prototype

Background: Concurrent Think Aloud Protocols

- Have been used for usability testing
- Ask users to verbalize what they are thinking while completing tasks
  - to gain critical insights from the information retained in their short-term memory (STM)
- The difficulty is to speak <u>continuously</u>
  - if users keep silent for a while, significant information may not be tracked down from STM

Approaches to Continuous Verbalization: Role of Facilitator

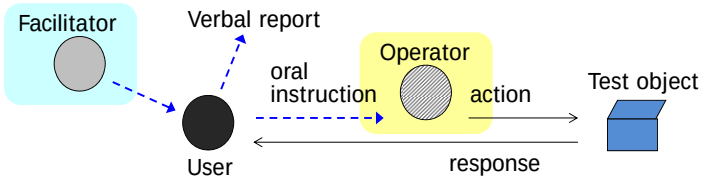


(1) Conventional method
- remind to keep talking (minimum intervention)

(2) Dialogue approach [Boren *et al*. 2000]
- use acknowledge tokens (*e.g., "OK" "yeah"*) continuously (proactive intervention)

may interrupt users' manner/pace of thinking, and given affirmative intension to the users

Our Approach: Indirect Oral Operation Method



- Facilitator only reminds users to keep talking
  - with minimum intervention
- Users are asked to
  - verbalize their thoughts (as usual), and
  - <u>speak every action to the operator</u> (w/o any manual op.)
- Allow users to
  - have more opportunity of verbalization
  - concentrate on the evaluation of a test object, even if the operational performance of a prototype is insufficient

Comparative Evaluation

- Compare the two operation methods
  - conventional manual operation
  - indirect oral operation
- Research questions
  - Do the two methods differ in terms of
    - the easiness of operation with oral operation
    - the number of utterances collected

Method

- 32 undergraduates without prior experience of think aloud protocols
  - Randomly divided into two groups of 16 participants each
  - One group for manual operation, and the other for oral operation
- All the participants were asked to work with two test objects (two tasks for each)
  - a prototype of a touch-screen digital camera
  - a working product of photo album software (to be used with mouse/keyboard UI)

Results: Ease of Operation

- Participants' rating was collected for the easiness of
  - (Q1) finding objects
  - (Q2) applying actions
  - Seven-point Likert scale (higher means more positive)
- Two-way ANOVA (operation methods, UIs)
  - Significant interactions revealed for both Q1 and Q2 (respectively, $p < 0.05$)
  - Simple main effect tests as follows …

Results: Ease of Operation (cnt'd)

- When the oral operation is used
  - mouse/keyboard (5.63) is easier than touch-panel screen (3.31)
    - To find a target object on a screen
  - no difference in the perceived easiness
    - To apply an action to the target object, which is already identified on the screen
- Intrinsic difficulty of the oral operation
  - in the process of identifying a target object

Results: Total Number of Utterances

Total number of utterances made for the four tasks

Verbalization Category		Manual Op. (n=16)		Oral Op. (n=16)	Welch's t-test
Explanation	(prediction)	7.81	<	17.19 *	$p < 0.05$
Procedure	(action)	0.25	<	71.00	—
Observation	(of results)	4.31	<	25.88 **	$p < 0.01$
Other		43.06	<	59.81	*n.s.*

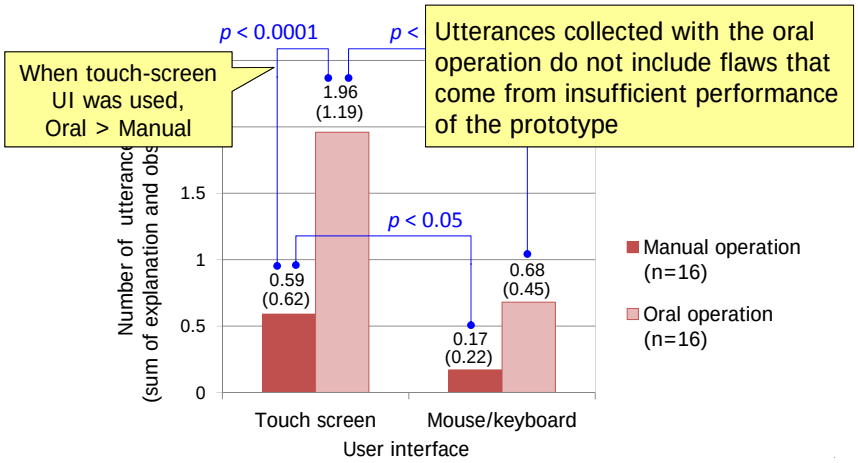
- When the oral operation was used
  - Numbers of utterances for the explanation and observation were significantly increased
  - These two categories would be important sources of discovering usability problems

Results: Utterances for Explanation and Observation

- Sum of explanation and observation utterances was analyzed
  - normalized as a value per task step
- A two-way ANOVA
  - Operation method (manual, oral)
  - UI (touch screen, mouse/keyboard)
- Significant interaction was observed
  - $F(1,30) = 10.2$, $p < 0.005$

Number of Utterances for Explanation and Observation

- Simple main effect tests



Concluding Remarks

- Oral operation method
  - will contribute to the increase of utterances for explanation and observation
    - depends on types of UI
- Further Study
  - Conduct more comparative evaluation
    - investigate the types of problems detected by the proposed method