LMU LUDWIG-MAXIMILIANS-UNIVERSITÄT MÜNCHEN

Carnegie Mellon

LFE Medieninformatik • Andreas Möller

Abschlussbericht Projektarbeit

Case Study for Supporting Intelligibility in Context-Aware Systems: IM Auto-Status Explanations

Verantw. Hochschullehrer: Prof. Dr. Heinrich Hußmann
Betreuerin: Sara Streng
Externer Betreuer: Prof. Anind Dey, Ph.D. (CMU)

30. Juni 2009

Andreas Möller
moellera@ifi.lmu.de

Outline

- Motivation
- Goals of this Work
- Realization of the Software
- User Study
- Conclusion and Outlook

Andreas Möller
moellera@ifi.lmu.de

Instant Messaging

- Important role in work environments
- Short interaction
  - Quick questions and clarifications
  - Coordination and scheduling
  - Impromptu social meetings
  - Contact with friends and family
- Longer interaction
  - Discussion of complex topics
  - Collaboration

- **But also:**
  - Receiver is disrupted by incoming messages
  - Sender has to wait for response
  - Negative effect on
    - Task performance
    - Memory
    - User‘s emotional state
  - Cause of annoyance

Andreas Möller
moellera@ifi.lmu.de

Possible Solutions

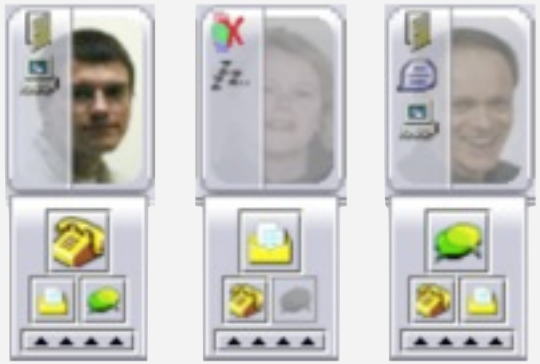
For the receiver

- Analysis of incoming messages
- Notification according to importance
- Reduction of unwanted work interruptions
- Example: *QnA*

[1]

For the sender

- Consideration of receiver's context
- Information about the receiver's presence and availability
- Reduction of waiting times due to unresponsiveness
- Example: *MyVine*



[6]

Andreas Möller
moellera@ifi.lmu.de

Intelligent Systems and Intelligibility

- Intelligent systems can never be perfect
- User acceptance depends on trust in the system
- Credibility can easily be lost!
  - When system does not behave as the user expects
  - Small errors can have relatively big impact
- Principles to support intelligibility
  - Information about system‘s understandings
  - Feedback (feedforward and confirmation)
  - Identity and action disclosure
  - Control
- Promising approach: Increasing intelligibility by explanations
  - Used in product recommendation system (Pu & Chen)

Andreas Möller
moellera@ifi.lmu.de

Outline

- Motivation
- Goals of this Work
- Realization of the Software
- User Study
- Conclusion and Outlook

Andreas Möller
moellera@ifi.lmu.de

Goals of this Work

- **„How can a context-aware system help the user to understand how it works?“**
- Design and implementation of a software that
  - provides responsiveness information about IM users
  - is able to explain its predictions
- Design of a long-term study which
  - allows to evaluate different versions of explanation provision
  - can provide findings about understanding and usefulness

Andreas Möller
moellera@ifi.lmu.de

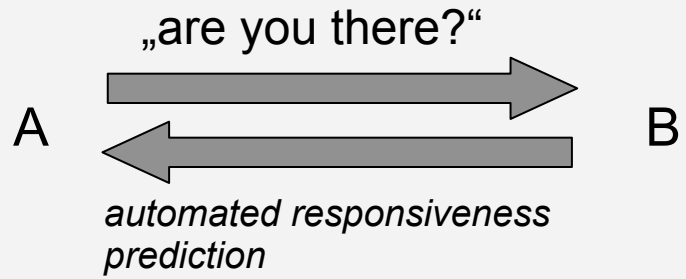
Outline

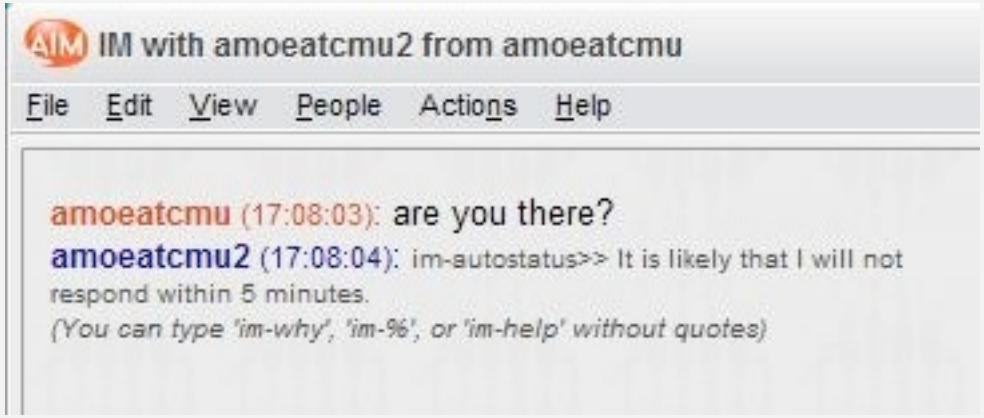
- Motivation
- Goals of this Work
- Realization of the Software
- User Study
- Conclusion and Outlook

Andreas Möller
moellera@ifi.lmu.de

IM Auto-Status Plugin

- Plugin for AOL Instant Messenger (AIM)
- „Auto-Status“ capability
- Principle of work:
  - A sends B a message
  - B‘s plugin predicts when B is likely to respond
  - The prediction is sent back to A
- User interface
  - Makes use of messaging window
  - No additional UI elements





Andreas Möller
moellera@ifi.lmu.de

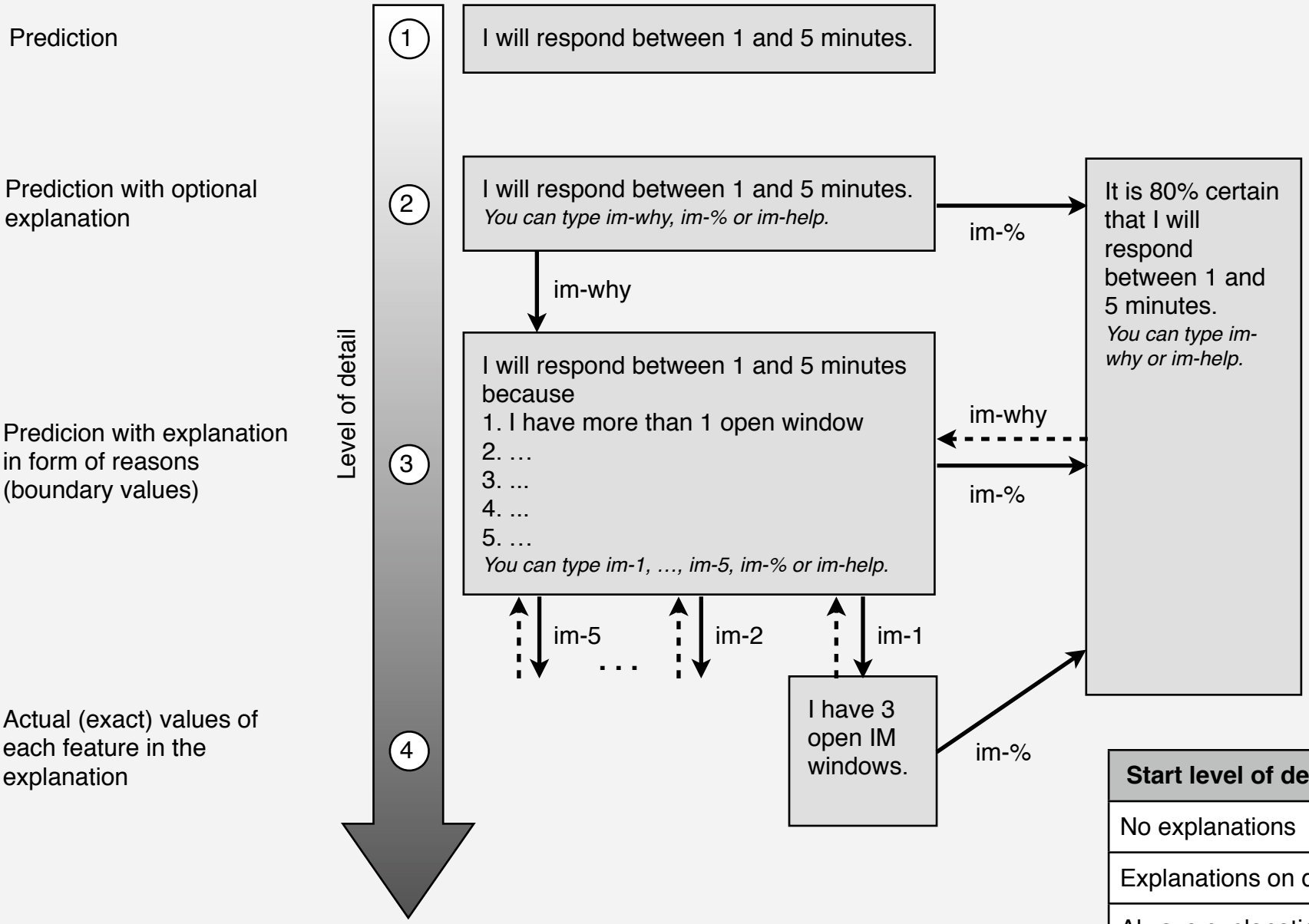
Explanation feature

- 5 reasons for why the system predicted like it did
- Additional features:
  - Prediction certainty
  - Exact values for reasons (will be explained later)
- Interaction using textual commands
- Three different versions of the plugin showing
  - Explanations for each prediction
  - Explanations on demand
  - No explanations

LMU LUDWIG-MAXIMILIANS-UNIVERSITÄT MÜNCHEN

Intelligibility in Context-Aware Systems: IM Auto-Status Explanations

Andreas Möller
moellera@ifi.lmu.de



Start level of detail in different versions	
No explanations	1
Explanations on demand	2
Always explanations	3

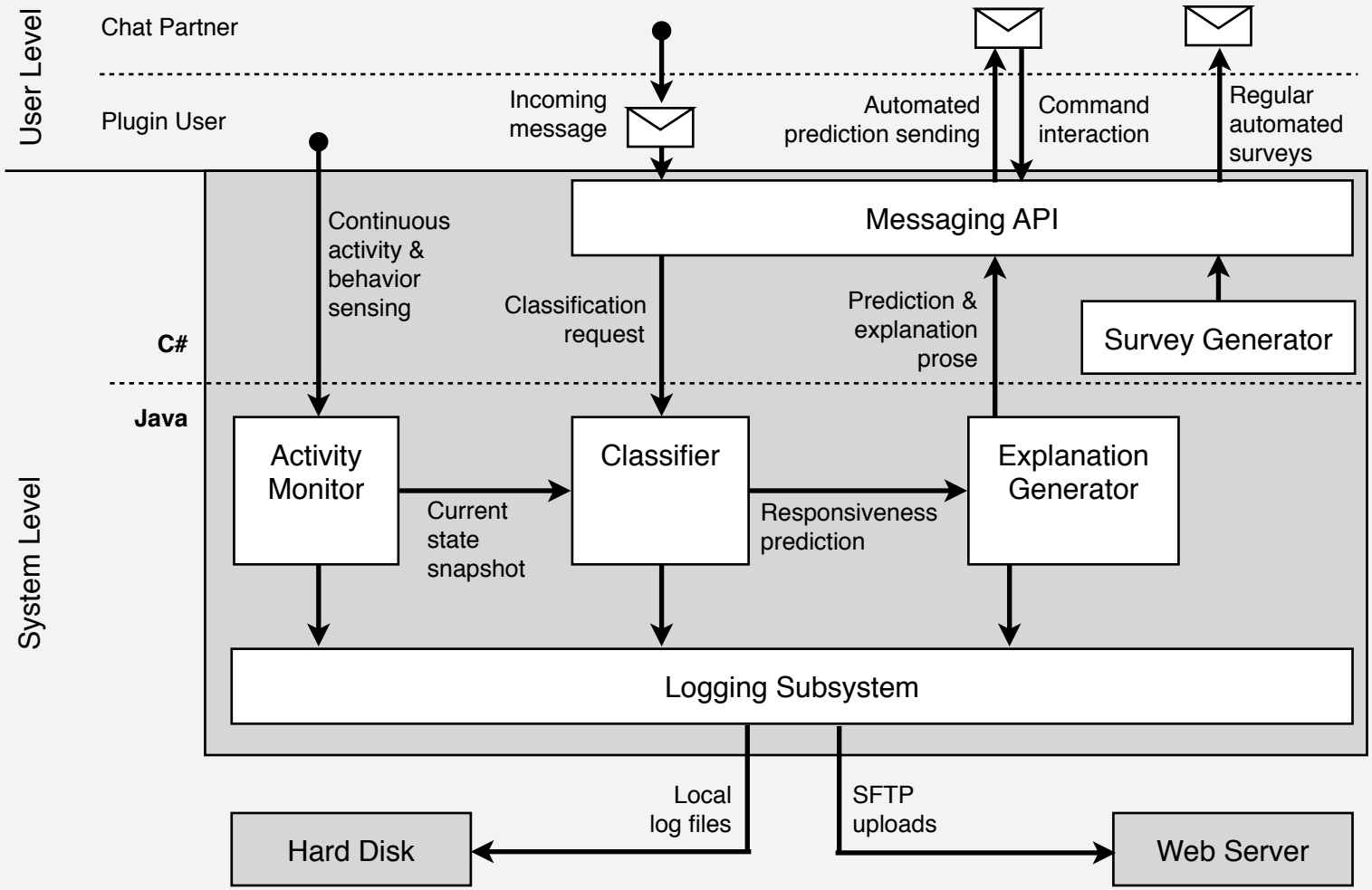
Andreas Möller
moellera@ifi.lmu.de

How does it work?

- Permanent activity monitoring
  - Desktop events: typing, switching between applications, ...
  - IM-related: incoming messages, time until IMs are answered, ...
- Responsiveness prediction using machine learning
  - Snapshot creation („instance“) using 68 features
  - Classification of user‘s current state using a decision tree
  - Pre-trained model
- Explanation generation
  - Extraction of the 5 most important features
  - Transformation in prose English

Andreas Möller
moellera@ifi.lmu.de

Software Architecture: Schematic Overview



Andreas Möller
moellera@ifi.lmu.de

Outline

- Motivation
- Goals of this Work
- Realization of the Software
- User Study
- Conclusion and Outlook

Andreas Möller
moellera@ifi.lmu.de

Hypotheses

- Recap: Three versions of explanation provision:
  - Always
  - On demand
  - Not at all (predictions only)
- **H1**: Users who get explanations find the application more useful and are more satisfied when using it.
- **H2**: Compared to users of the on-demand version, users who always get explanations will understand the application better. Yet, they will find it more obtrusive than users of the on-demand version.
- **H3**: Subjects using the low accuracy version find the explanations more useful, and improve more than those using the high accuracy version.

Andreas Möller
moellera@ifi.lmu.de

Study Setup

- 12 groups, each consisting of at least 3 participants
- One person in each group installs the plugin (*receiver*), the others receive predictions about him/her (*senders*)
- Daily chats during 2 phases of 3 weeks each
- Two different explanation types in phase 1 and 2

Group	Prediction Accuracy	Explanation type	
		Phase 1	Phase 2
1	High	No explanations	Always On
2		No explanations	Always On
3		No explanations	On Demand
4		Always On	No explanations
5		Always On	No explanations
6		On Demand	No explanations
7	Low	No explanations	Always On
8		No explanations	Always On
9		No explanations	On Demand
10		Always On	No explanations
11		Always On	No explanations
12		On Demand	No explanations

Andreas Möller
moellera@ifi.lmu.de

Data Collection

- Regular surveys
  - Daily questions, referring to particular IMs
    - Agreement with prediction and user predictions
    - User understanding
  - Weekly online questionnaires, referring to the overall impression
    - User understanding
    - Perception of usefulness, reliability, and trust
    - Changes in interaction with buddies (qualitative)
- Personal interview (phone or IM)
- Logging by the plugin
  - Actual correctness of the predictions
  - Timing/behavior pattern changes

Andreas Möller
moellera@ifi.lmu.de

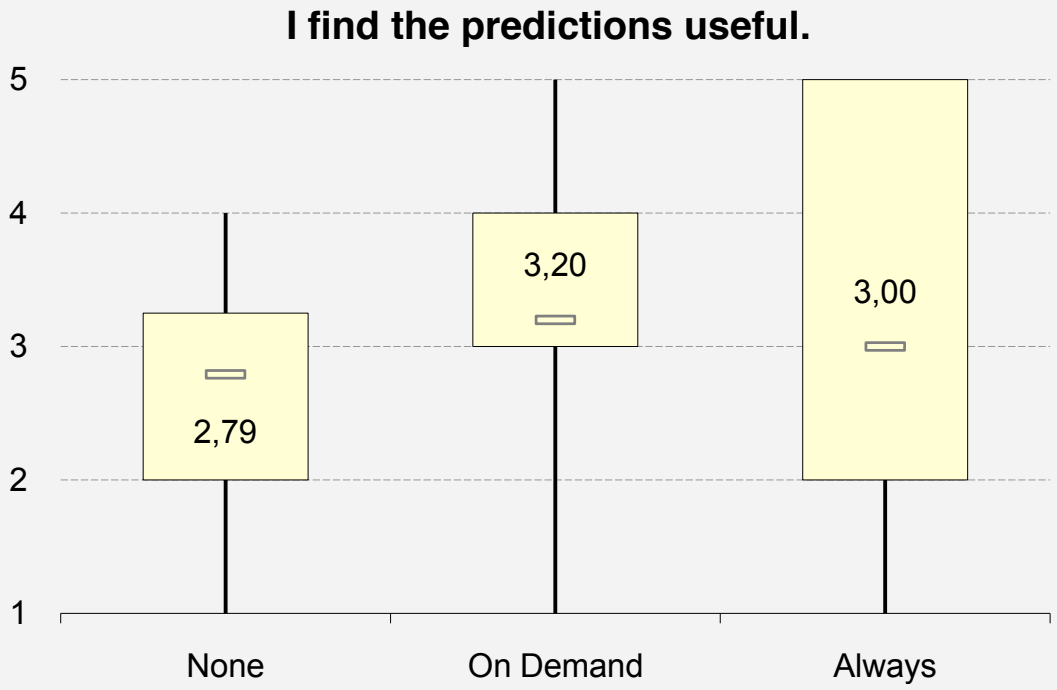
The study is not yet completed.

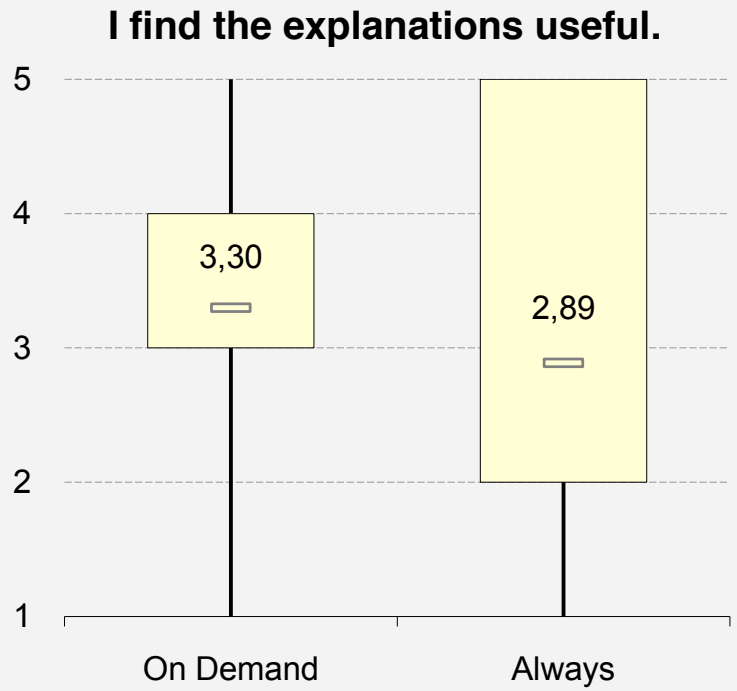
- Why?
  - Recruitment
    - Less interested parties than expected → extension to other cities
  - Subscription phase
    - Incomplete/bogus subscriptions
    - Groups not meeting all requirements
  - Decline of agreement terms („daily chats“)
  - Technical issues
- Difficulties in extensive long-term studies like this:
  - High effort for participants
  - Group subscriptions
  - Reduced commitment due to missing face-to-face relationship

Andreas Möller
moellera@ifi.lmu.de

First Trends (1)

- Usefulness of predictions and explanations





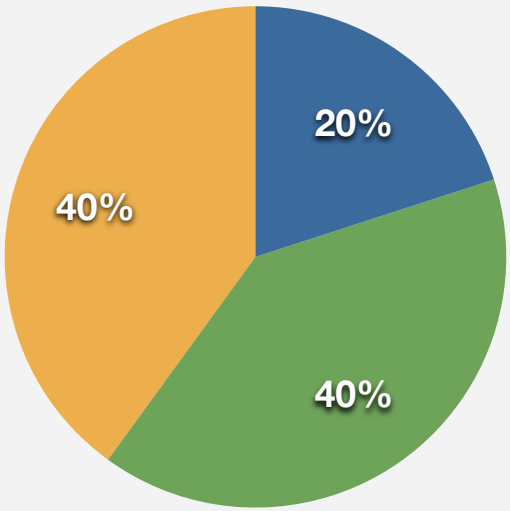
1: strongly disagree, 5: strongly agree

Andreas Möller
moellera@ifi.lmu.de

First Trends (2)

- Preference for respective other version

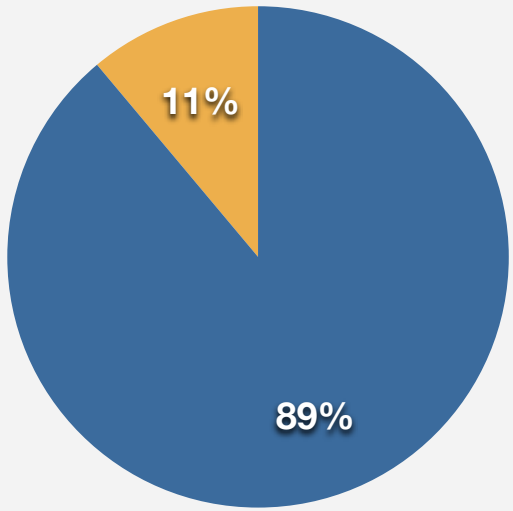
Would you rather have always explanations?



Yes No Not Sure

among subjects who get explanations on demand

Would you rather have explanations on demand?



Yes No Not Sure

among subjects who get always explanations

Andreas Möller
moellera@ifi.lmu.de

First Trends: Summary

- Predictions are appreciated
- Usefulness increases with explanations
- Strong preference towards on-demand version
  - Usefulness of explanations considered higher
  - „Always explanations“ too obtrusive
  - No advantages for „always explanations“ shown (yet)
- H1 and H2 are likely to be verified

Andreas Möller
moellera@ifi.lmu.de

Outline

- Motivation
- Goals of this Work
- Realization of the Software
- User Study
- Conclusion and Outlook

Andreas Möller
moellera@ifi.lmu.de

Conclusion and Outlook

- Responsiveness predictions and explanations
  - Can help senders to estimate response time
  - Can help reduce work interruptions
- Future of this project
  - Relation between effect of explanations & error rate
  - Changes in understanding and behavior
- Outlook
  - Downloadable version at AIM Gallery?
  - Different approach: change salience of message notifications according to responsiveness

Andreas Möller
moellera@ifi.lmu.de

Thank you for your attention!

Questions and Discussion

Andreas Möller
moellera@ifi.lmu.de

References

[1] Daniel Avrahami and Scott E. Hudson. QnA: augmenting an instant messaging client to balance user responsiveness and performance. In CSCW '04: Proceedings of the 2004 ACM conference on Computer supported cooperative work, pages 515–518, New York, NY, USA, 2004. ACM.

[2] Brian P. Bailey, Joseph A. Konstan, and John V. Carlis. Measuring the effects of interruptions on task performance in the user interface. In IEEE Conference on Systems, Man, and Cybernetics 2000 (SMC 2000, pages 757–762. IEEE Press, 2000.

[3] Brian P. Bailey, Joseph A. Konstan, and John V. Carlis. The effects of interruptions on task performance, annoyance, and anxiety in the user interface. In Proceedings INTERACT '01, pages 593–601. IOS Press, 2001.

[4] Victoria Bellotti and Keith Edwards. Intelligibility and accountability: human considerations in context-aware systems. Hum.-Comput. Interact., 16(2):193–212, 2001.

[5] Edward B. Cutrell, Mary Czerwinski, and Eric Horvitz. Notification, disruption, and memory: Effects of messaging interruptions on memory and performance. In Proceedings INTERACT '01, pages 263–269. IOS Press, 2001.

[6] James Fogarty, Jennifer Lai, and Jim Christensen. Presence versus availability: the design and evaluation of a context-aware communication client. Int. J. Hum.-Comput. Stud., 61(3):299–317, 2004.

[7] B. J. Fogg and Hsiang Tseng. The elements of computer credibility. In CHI '99: Proceedings of the SIGCHI conference on Human factors in computing systems, pages 80–87, New York, NY, USA, 1999. ACM.

[8] Ellen Isaacs, Alan Walendowski, Steve Whittaker, Diane J. Schiano, and Candace Kamm. The character, functions, and styles of instant messaging in the workplace. In CSCW '02: Proceedings of the 2002 ACM conference on Computer supported cooperative work, pages 11–20, New York, NY, USA, 2002. ACM.

[9] Scott LeeTiernan, Edward Cutrell, Mary Czerwinski, and Hunter Hoffman. Effective notification systems depend on user trust. In Proceedings of Human-Computer Interaction— Interact, pages 684–685, Tokyo, Japan, July 2001. IOS Press.

[10] Bonnie A. Nardi, Steve Whittaker, and Erin Bradner. Interaction and outeraction: instant messaging in action. In CSCW '00: Proceedings of the 2000 ACM conference on Computer supported cooperative work, pages 79–88, New York, NY, USA, 2000. ACM.

[11] Pearl Pu and Li Chen. Trust building with explanation interfaces. In IUI '06: Proceedings of the 11th international conference on Intelligent user interfaces, pages 93–100, New York, NY, USA, 2006. ACM.