

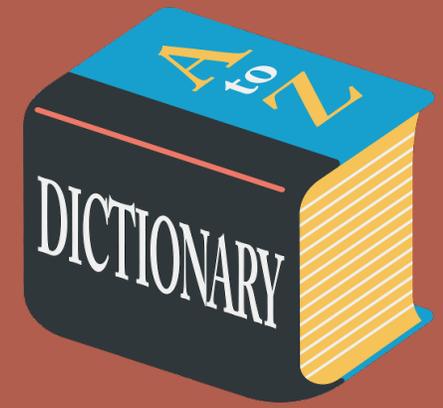


Transparency and Responsibility Attribution in Smart Product Failure

Nancy Xia & Christos Bechlivanidis

Why?

- Smart products proliferate; 20bn IoT devices in the world (Statista, 2020)
- Their complexity and interoperability increases
- Failure is inevitable
- How do consumers attribute responsibility and what information do they need?



Defining the terms

- What are the properties of “smart” products?
 - Connectivity, ability to learn, reactivity, autonomous decision-making... (Rijsdijk et al., 2007)
- What do we mean by transparency?
 - “Actions, processes and data are made open to inspection by publishing information about the project in a complete, open, understandable, easily-accessible and free format.” Data Ethics Framework (2020)
- What constitutes an explanation?
- How do people attribute responsibility?

Delegation and responsibility

- When delegating a task, we commonly delegate responsibility.
- But blame attribution requires not only that the agent is causally critical but also that they are capable and can foresee the outcomes of their actions.



What is transparency?

- What information do the end users need to know?
 - Statistical information, e.g. reliability and past performance (Dzindolet et al. , 2003)
 - What needs explaining?
 - Normal vs unexpected behaviour
 - How and when should explanations be given?

Aims

- Who do people blame when smart devices fail?
- How are blame attributions influenced by (types of) explanations?



Current project

- Present scenarios of failure of (novel) smart product to N=135 participants.
- Varying the presence, type and timing of explanations
- Ask participants to assign responsibility to the various involved parties:
 - The user
 - The smart product
 - The manufacturer
 - The retailer
 - The regulator

Current project

- Each participant sees descriptions and usage instructions for 4 imaginary smart products: smart jacket, smart ladder, smart oven, and smart iron
- Then they read a scenario where the smart product fails with minor consequences, e.g. the smart oven burnt the turkey, the smart iron left a scorch mark on your shirt etc

Current project

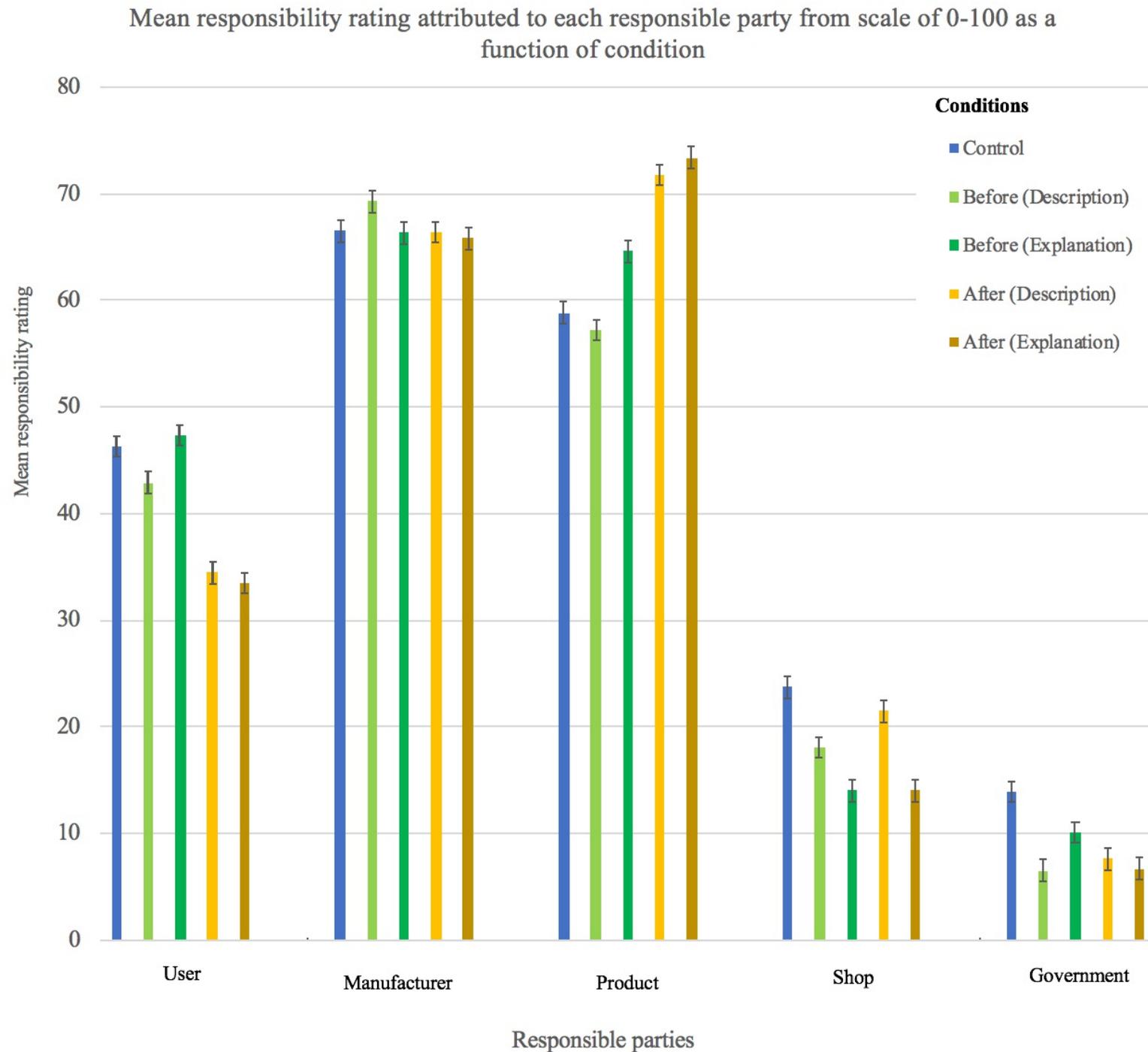
- Critically, some participants saw no explanation about the product behaviour, others saw either a **description** or an **explanation** of its behaviour, either **before** or **after** the failure

Description		Explanation	
Before	After	Before	After
"The turkey will be cooked at 180 degrees for 315 minutes"	"The turkey was cooked at 180 degrees for 315 minutes"	"The turkey will be cooked at 180 degrees for 315 minutes because it is heavier than average"	"The turkey was cooked at 180 degrees for 315 minutes because it was heavier than average"

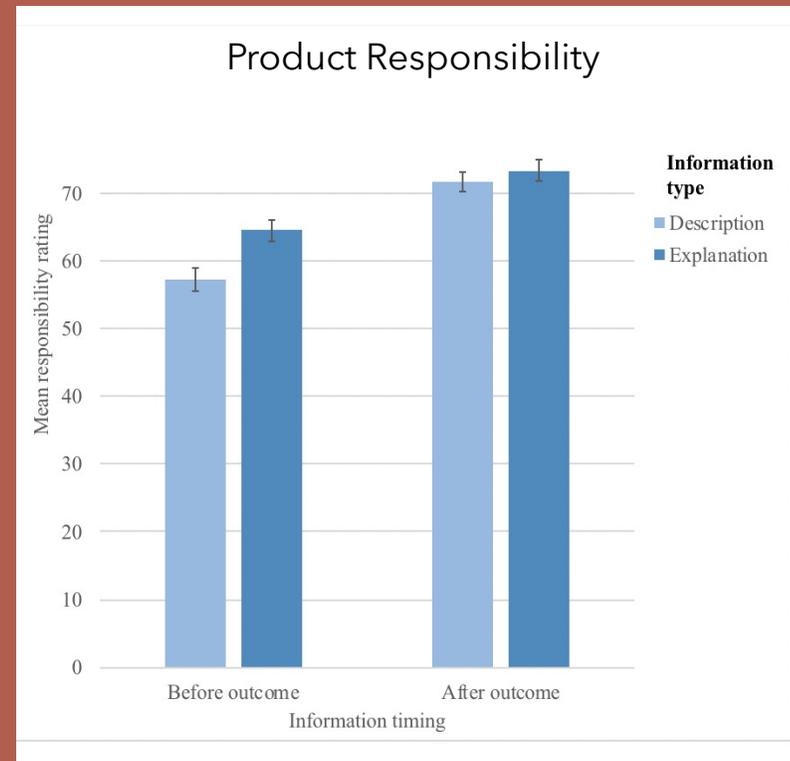
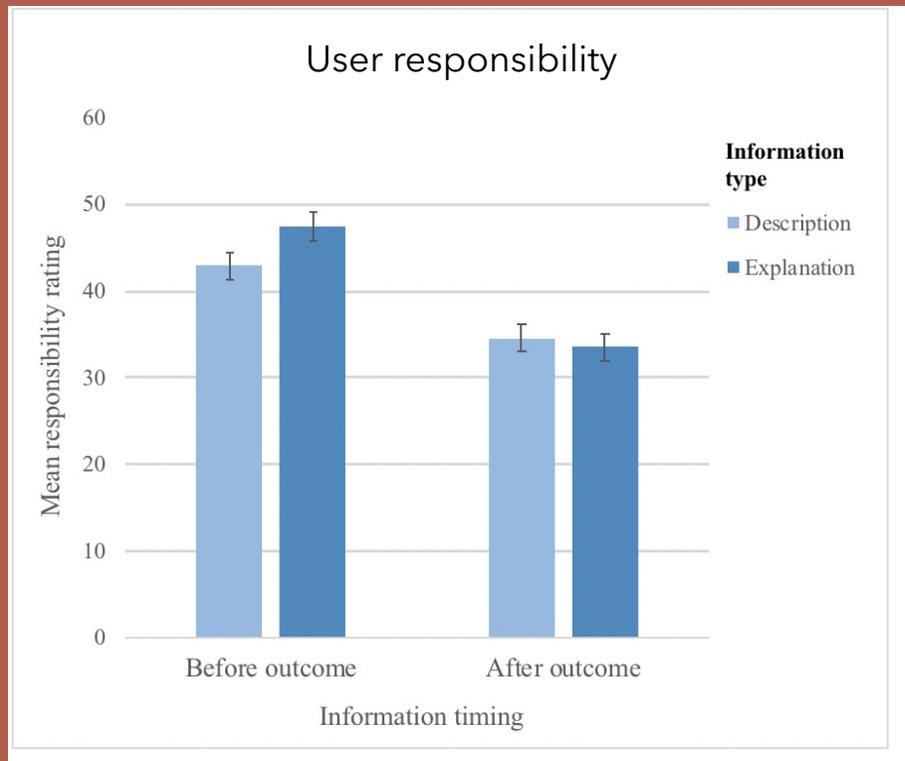


Results

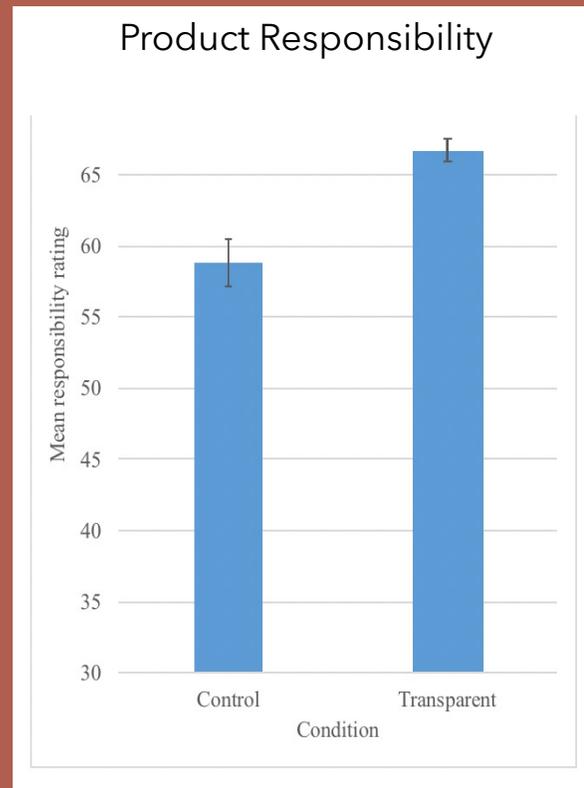
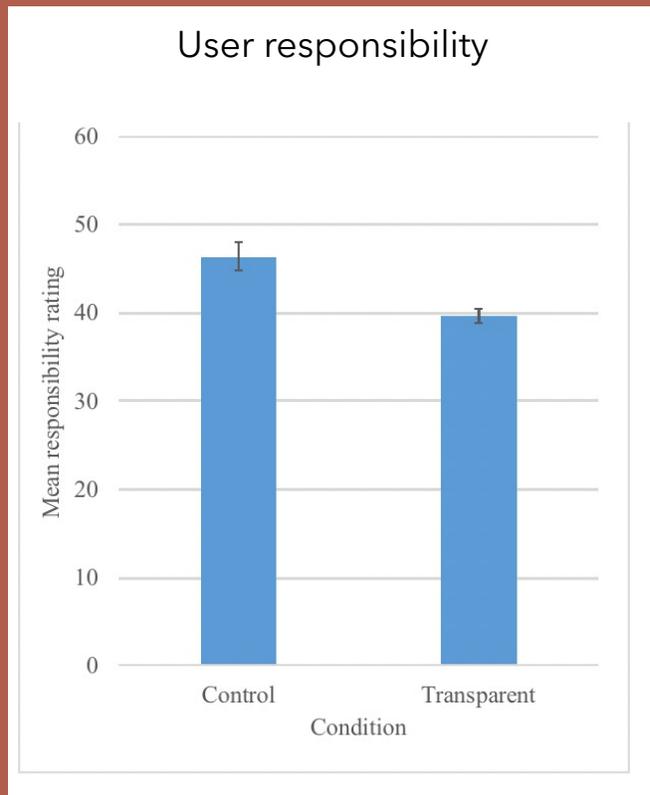
- Overall people blame themselves less than the manufacturer or the product itself
- Some interesting differences between product/manufacturer blame



Timing of information



Presence of information



The more transparent the product behaviour, the more participants blame the product and the less they blame themselves.

Transparency paradox?

- We expected that the presence of transparency would lead people to attribute less responsibility to the product and more to the user but we find the opposite pattern
- Similar findings in Kim & Hinds (2006) in scenarios of human-robot cooperation.
- Current hypothesis: Explanation is revealing capability which increases blame

Next steps

- Replicate the study
- Distinguish between the timing of information and the opportunity to act based on information
- Distinguish between information that reveals the capacities of the device and explanations of its current behaviour



Thanks!