Context for Actions:
The time course of action-scene compatibility

Oleg Vrabie, Marius Zimmermann, & Angelika Lingnau

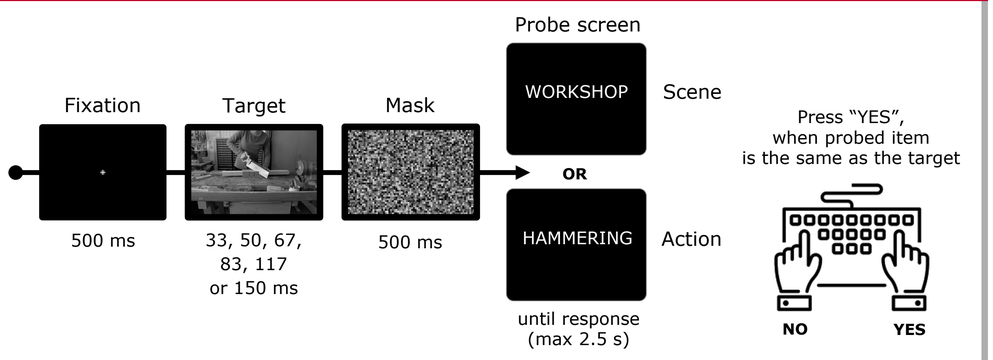
Chair of Cognitive Neuroscience, Institute of Psychology, University of Regensburg, Germany

Introduction

- Contextual associations between actions, objects and scenes have been suggested to be exploited for action understanding [1]
- Scenes are known to facilitate both object and action recognition, especially when difficult to recognize [2, 3]
- Incompatible scenes have been shown to slow down action recognition [3, 4]
- Objects in consistent backgrounds and backgrounds with consistent objects are reported more accurately [5]
- It has been proposed that initial coarse information triggers, in parallel, multiple hypotheses about objects as well as a global context, the latter quickly resolving competition among object hypotheses [6]

❑ How does **action-scene compatibility** affect each other's recognition?
- ❑ Are **actions profiting** from compatible scenes?
- ❑ Are **scenes profiting** from compatible actions?
- ❑ How do **compatibility** effects depend on **exposure duration**?

Methods

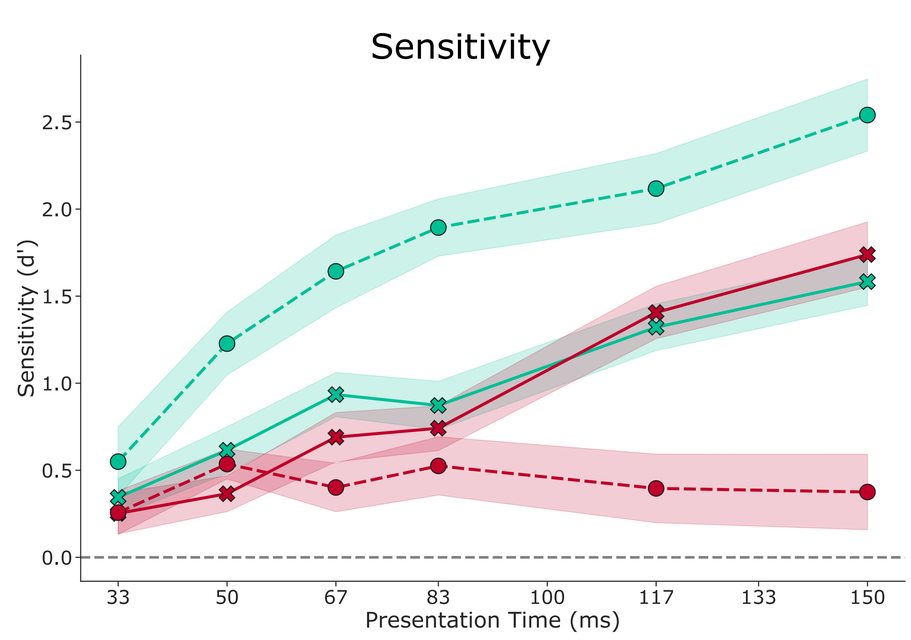


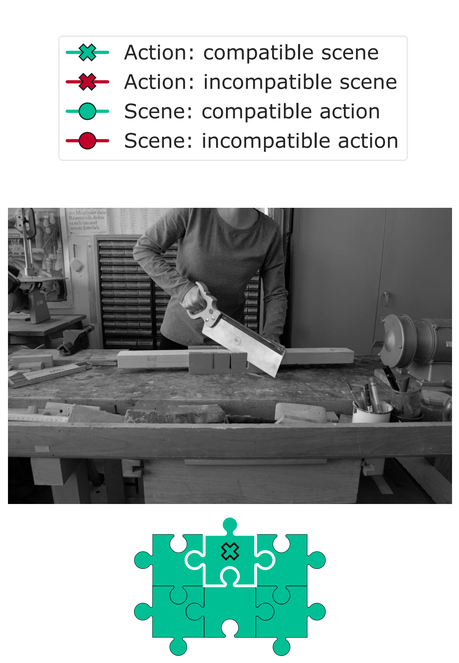
- Participants: N = 20 | right-handed | native German speakers
- Within-subject factors: *Presentation Time* | *Compatibility* | *Probe Type*

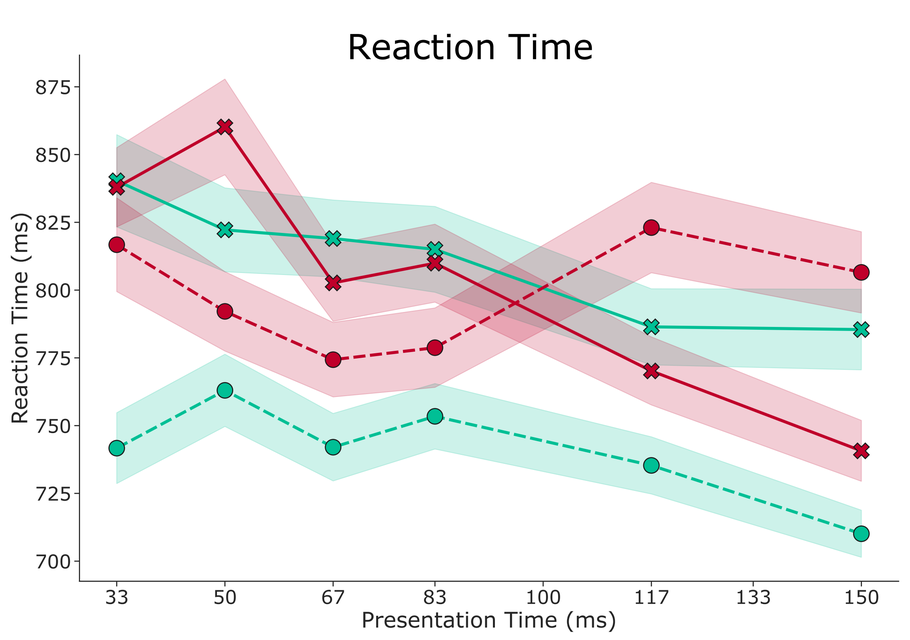
	Compatibility	Scene		
		Kitchen	**Office**	**Workshop**
Action	Compatible	Cutting	Hole-punching	Hammering
		Grating	Stamping	Painting
		Whisking	Stapling	Sawing
	Incompatible	Hole-punching	Cutting	Stapling
		Stamping	Grating	Hole-punching
		Stapling	Whisking	Stamping
		Hammering	Hammering	Cutting
		Painting	Painting	Grating
		Sawing	Sawing	Whisking

**Actions selected based on criteria described in Wurm & Schubotz [4].*

Results







- Both actions and scenes were recognized better with longer exposure
  - Three-way ANOVA, Probe Type ✻ Compatibility ✻ Presentation Time: $F(5, 90) = 8.4$, $p < .001$
  - Two-way ANOVA (Actions), Presentation Time: $F(5, 90) = 32.9$, $p < .001$
  - Two-way ANOVA (Scenes), Presentation Time: $F(5, 90) = 18.5$, $p < .001$
- Sensitivity of scene recognition was strongly modulated by action compatibility, increasingly so with longer exposure
  - Two-way ANOVA (Scenes)
    - Compatibility: $F(1, 18) = 53.0$, $p < .001$
    - Compatibility ✻ Presentation Time: $F(5, 90) = 9.6$, $p < .001$
- Sensitivity of action recognition was not modulated by scene compatibility
  - Two-way ANOVA (Actions), Compatibility: $F(1, 18) = 1.6$, $p > 0.2$

- Scenes were reported faster than actions
  - Three-way ANOVA
    - Probe Type ✻ Compatibility ✻ Presentation Time: $F(5, 90) = 2.7$, $p < .05$
    - Probe Type: $F(1, 18) = 10.5$, $p < .01$; $M_{Action} = 807 \pm 20$ ms; $M_{Scene} = 768 \pm 16$ ms
- Actions were reported faster with longer exposure
  - Two-way ANOVA (Actions), Presentation Time: $F(5, 90) = 7.8$, $p < .001$
- Scenes were reported faster when presented with compatible actions
  - Two-way ANOVA (Scenes), Compatibility: $F(1, 18) = 23.1$, $p < .001$

Discussion

- We found no evidence supporting the view that action recognition benefits from compatible scenes, in contrast to previous studies [3, 4]
- Scene recognition is modulated by action compatibility, increasingly so with longer exposures, with reduced performance when presented with incompatible actions
- As it has been proposed for objects in scenes [6], we assume that actions are processed in parallel and complementarily to scenes
- Parallel to a global context, initial coarse scene information triggers expectations about potential action hypotheses, biasing the competition between the action candidates
- When an action does not match the scene, we hypothesize that the scene evidence is suppressed, leading to a delayed response and a lower accuracy selectively for scenes

References

[1] Wurm et al. (2012) The Context–Object–Manipulation Triad. *J Cogn Neurosci*
[2] Bradnman & Peelen (2017) Interaction between scene and object processing. *J Neurosci*
[3] Wurm & Schubotz (2017) What's she doing in the kitchen? *Psychon Bull Rev*
[4] Wurm & Schubotz (2012) Squeezing lemons in the bathroom. *Neuroimage*
[5] Davenport & Potter (2004) Semantic Consistency in Perception. *Psychological Science*
[6] Trapp & Bar (2015) Prediction, context, and competition in visual recognition. *Ann. N.Y. Acad. Sci.*

oleg.vrabie@ur.de @oleg_vrabie oleg-vrabie

This research was supported by funds from German Research Foundation (DFG), reference number LI 2840/2-1 and LI 2840/4-1 awarded to A.L.