

CREATING MOUSE-TRACKING EXPERIMENTS AND ANALYZING MOUSE-TRACKING DATA

Pascal Kieslich (University of Mannheim) & Dirk Wulff (University of Basel) Workshop at the EADM Summer School 2018 in Salzburg, Austria

Workshop agenda

Mouse-tracking introduction (Monday)

- General introduction
- Your task
- Develop & present experimental design

Creating mouse-tracking experiments (Tuesday)

- Introduction to OpenSesame & mousetrap-os plugin
- Build & preregister experiment
- Run experiments

Analyzing mouse-tracking data (Wednesday)

- Introduction to R & mousetrap package
- Covering both basic and advanced analyses and visualizations
- Analyze your data

Preparations (before the workshop)

- Read book chapter by <u>Kieslich et al. (in press)</u>
- Outline two example experiments in your group (meeting the outlined requirements) and describe them in a paragraph
- Upload your ideas in one file name 'GroupX.doc' onto OSF (<u>Project Ideas</u>)



DAY 1: MOUSE-TRACKING INTRODUCTION

Pascal Kieslich (University of Mannheim)
Workshop at the EADM Summer School 2018 in Salzburg, Austria

Mouse-tracking introduction (Monday)

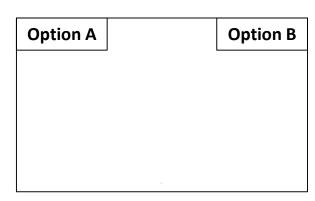
- □ 13:00-14:30 General introduction to mouse-tracking
 - Paradigm and assumptions
 - Implementation and analysis
 - Previous applications
- □ 14:30-15:00 Introduction to task
 - Type of experiments considered
 - Your tasks during the workshop
- 15:00-17:00 Develop experimental design conceptually
- 17:00-18:00 Present experimental design in plenum



Paradigm & assumptions

Paradigm & assumptions

- Mouse-tracking (aka. response dynamics)
 - **□** Continuous recording of mouse movements
 - while participants decide between different spatially separated options on a screen

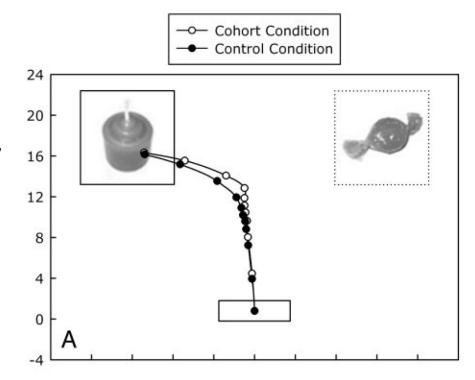


Assumptions

- Cognitive processing continuously revealed in motor responses (Spivey & Dale, 2006)
- "Hand in motion reveals mind in motion" (Freeman et al., 2011)
- Mouse movements reveal tentative commitments to and conflict between choice options during decision process

Seminal article by Spivey et al. (2005)

- Study on spoken word recognition
 - Instruction: "Click the candle"
- Spatial attraction of hand movement
 - Greater towards phonologically similar distractor ("candy")
 - Than towards phonologically dissimilar distractor ("dice")
- Evidence
 - Suggests parallel processing of auditory input activating competing representations



Main applications

- Mouse-tracking allows for testing psychological theories
- Two major applications (cf. review by Stillman et al., 2018)
 - Provides fine-grained measure for amount of conflict between response options
 - → test predictions about which factors (contextual factors, individual differences) influence amount of conflict for specific decision
 - Assess temporal development of conflict and its resolution
 → test models that make predictions how decisions unfold over time
 (e.g., decide between single vs. dual process models)

Application domains

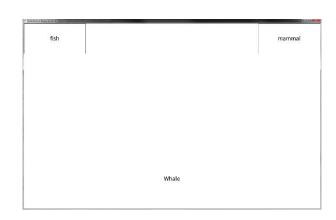
- Application of mouse-tracking in a growing number of psychological domains (Reviews by Freeman, in press; Stillman et al., 2018)
 - Semantic processing (e.g., Spivey et al., 2005; Dale & Duran, 2011)
 - Social cognition (e.g., Freeman et al., 2008; Freeman & Ambady, 2011)
 - Learning and memory (e.g., Dale et al., 2008; Koop & Criss, 2016)
 - Self-control (e.g., Sullivan et al., 2015; Stillman et al., 2017)
- In the last years also extended to JDM research
 - Intertemporal choice (Dshemuchadse et al., 2013)
 - Moral dilemmas (Koop, 2013)
 - Decisions under risk (Koop & Johnson, 2013)
 - Social dilemmas (Kieslich & Hilbig, 2014)
 - Judgmental biases (Szaszi et al., 2018; Travers et al., 2016)



Replication study of Dale et al. (2007)

Animal categorization task

- Typical exemplars only share features with correct category (e.g., cat as mammal)
- Atypical exemplars share both features with correct and competing category (e.g., whale with mammal and fish)



Main hypothesis

- Increased competition when categorizing atypical exemplars
- Mouse trajectories with deviation towards competing category
- Replication study (Kieslich & Henninger, 2017)
 - Same material (translated into German) and procedure, but higher resolution and different aspect ratio
 - \sim N = 60 students from the University of Mannheim
 - Material, data, and analyses at https://github.com/pascalkieslich/mousetrap-resources

Methodological considerations

- General challenge when designing a mouse-tracking study
 - Movements should reflect developing commitment not information search (≠ eye-tracking or Mouselab)
 - → minimize amount of new information after tracking onset
 - Preferences should not develop before tracking starts
 - → critical information should only be made available at the last moment
- Mouse start positions should be comparable across trials
 - Participants have to click on a centered button to start the trial
 - Exactly identical start positions across trials achieved by resetting mouse or by computational alignment during analysis
- Counterbalancing positions across trials / participants
 - Vary which option is presented on which side (left vs. right)
 - Can be done between trials or between participants (depending on study)



Typical analyses steps

Calculate **Preprocess** Inspect Aggregate Compare Filter trials Trial-level Plot of Potential Comparison of

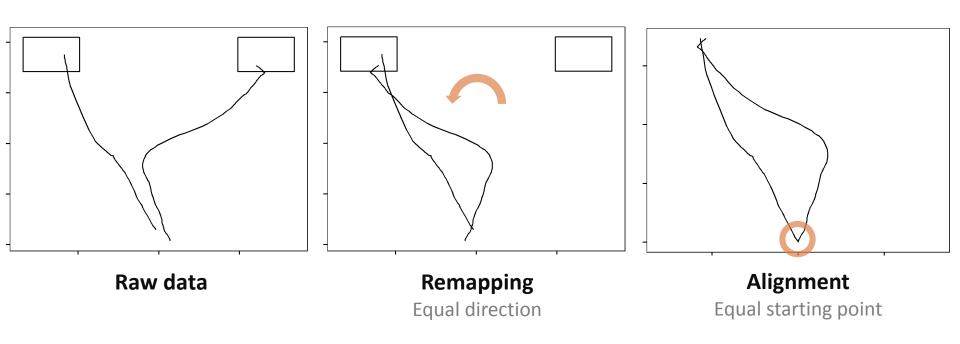
- Alignment of trajectories
- Remapping of trajectories
- Time (and space) normalization
- indices (e.g., MAD for curvature)
- Dynamic measures (e.g., development of acceleration over time)
- individual trajectories
- Distribution of trial-level indices
- Identification of chaotic trajectories
- aggregation of trajectories and trial measures
- Per participant and condition
- measures and trajectories
- Typically between experimental conditions in within designs

Analyses steps implemented in the mousetrap R package

More information: http://pascalkieslich.github.io/mousetrap/

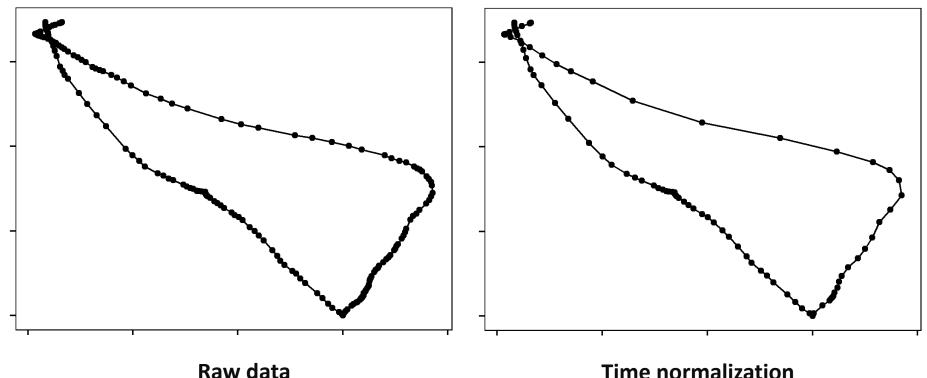
Available from CRAN: install.packages("mousetrap")

Data preparation: Remapping and alignment



Data preparation: Time normalization

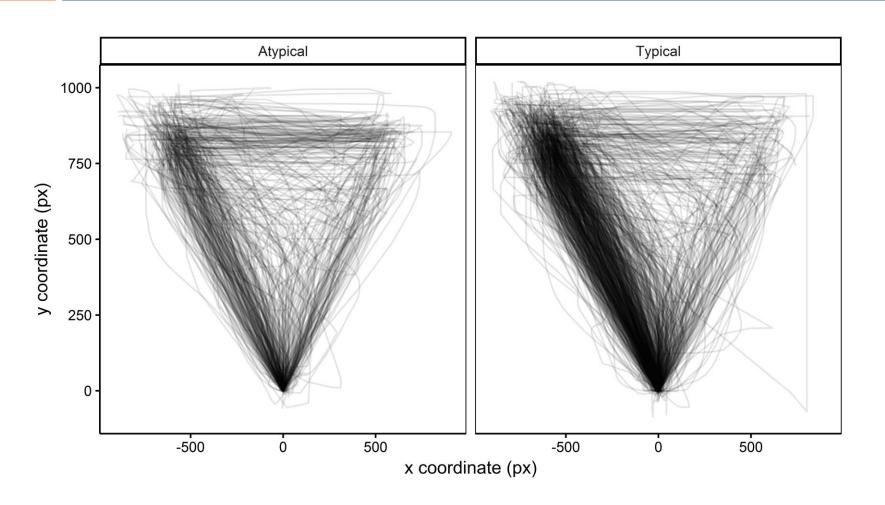
- □ Trials with differing response time vary regarding number of recorded coordinates
- To permit averaging across trials: time-normalization (cf. Spivey et al., 2005)
- Each trajectory divided into 101 equally spaced time steps using linear interpolation



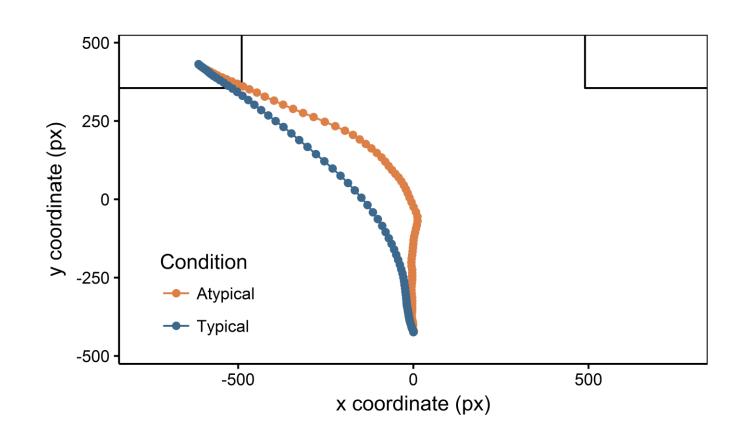
Constant sampling rate → Absolute time

Time normalization
Relative time steps

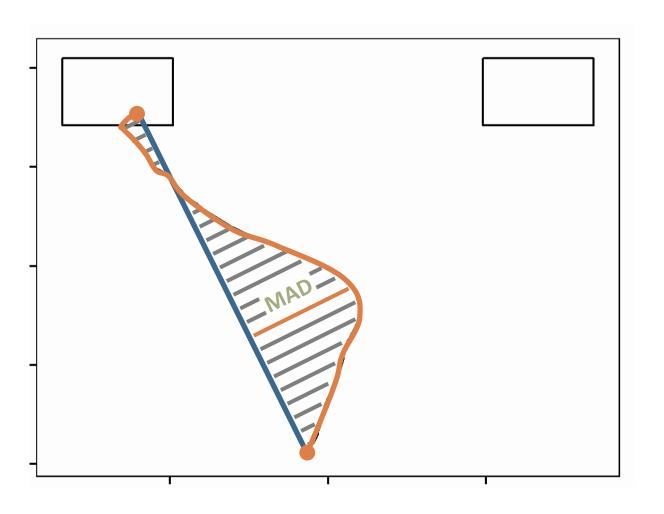
Time-normalized trajectories per condition



Average time-normalized trajectories



Selected measures for trajectory curvature



Measures of curvature quantify perpendicular distance between observed trajectory and an idealized straight line

- Maximum absolute
 deviation (MAD)
 McKinstry, Dale, & Spivey (2008)
- Average deviation (AD) Koop & Johnson (2011)
- Area under curve (AUC) Spivey, Grosjean, & Knoblich (2005)

Typical analyses steps

Preprocess Calculate Inspect Aggregate Compare

- Filter trials
- Alignment of trajectories
- Remapping of trajectories
- Time (and space) normalization
- Trial-level indices (e.g., MAD for curvature)
- Dynamic measures (e.g., development of acceleration over time)
- Plot of individual trajectories
- Distribution of trial-level indices
- Identification of chaotic trajectories
- Potential aggregation of trajectories and trial measures
- Per participant and condition
- Comparison of measures and trajectories
- Typically between experimental conditions in within designs

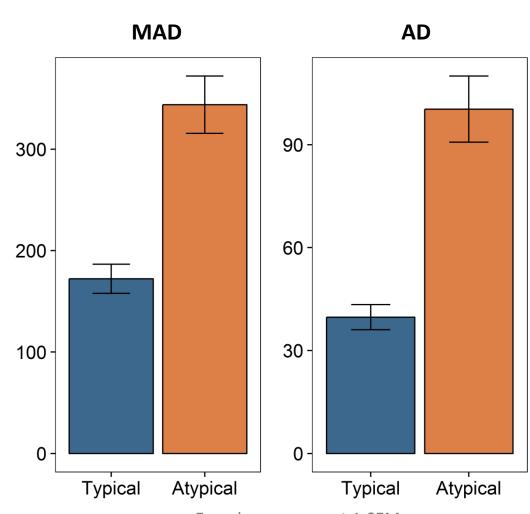
Analyses steps implemented in the mousetrap R package

More information: http://pascalkieslich.github.io/mousetrap/

Available from CRAN: install.packages("mousetrap")

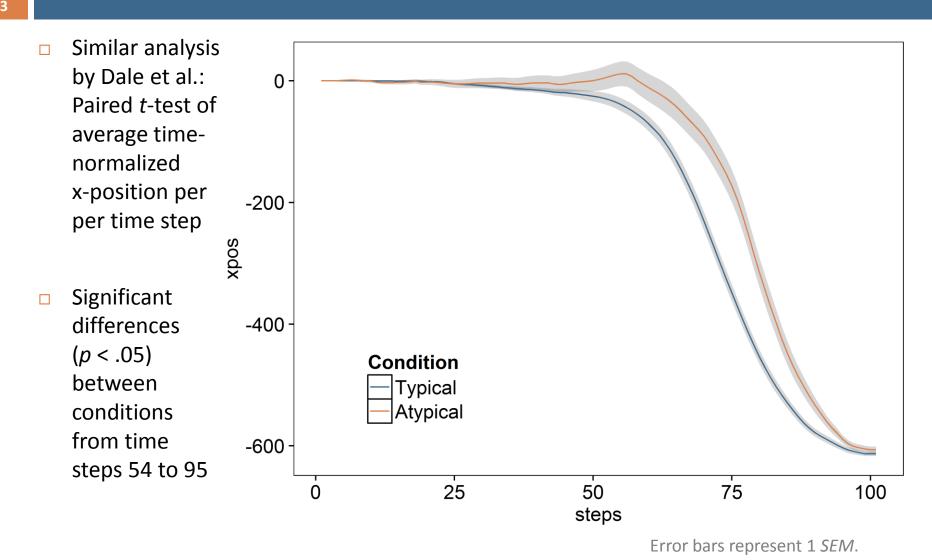
Comparison of (maximum) absolute deviations

- MAD larger for atypical exemplars
 - $d_z = 0.87, p < .001$
 - $BF_{10} = 1.57 * 10^6$
- AD larger for atypical exemplars
 - $d_7 = 0.87, p < .001$
 - $BF_{10} = 1.78 * 10^6$



Error bars represent 1 SEM.

Average x-positions per time step



Selected mouse-tracking measures

Measure	Definition	Possible interpretation	Example
Maximum absolute deviation (MAD)	Maximum deviation from idealized trajectory	Maximum attraction of non-chosen option	McKinstry et al. (2008)
Average Deviation (AD)	Mean deviation from idealized trajectory	Average attraction of non-chosen option	Koop & Johnson (2011)
Area under curve (AUC)	Geometric area between actual and idealized trajectory	Total attraction of non- chosen option	Spivey et al. (2005)
x-flips (xpos_flips)	Number of directional changes along x-axis	Instability, reversal of the momentary valence	Koop & Johnson (2013)
x-reversals (xpos_reversals)	Number of crossings of the y-axis	General reversal of preference	Koop & Johnson (2013)

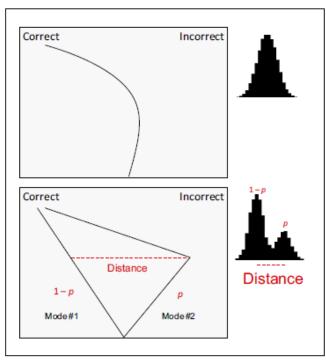
Analytical and theoretical challenges

- Interpretation of measures still needs to be validated
- Multitude of mouse-tracking measures available
 - Often highly correlated in practice
 - There is no standard yet which measure should be used
 - ensure that result does not depend on the specific measure used
 - → decide which is the measure of interest a priori / conduct preregistered replications of your findings
- Consider effects of aggregation by inspecting distribution of trajectories and indices on the trial level

Smooth competition vs. abrupt shifts

- □ Different assumptions about response process (e.g., Hehman et al., 2015)
 - Single process
 - "smooth graded competition" in all trials
 - Continuous competition between response options
 - Dual process
 - "abrupt shifts" / Change of Mind in some trials: Initial movement towards one option, then reversal and choice of other option
 - Straight movements in other trials
- Statistical analysis of AUC or MAD distribution (Freeman & Dale, 2013)
 - "smooth graded competition" → unimodal
 - "abrupt shifts"

 bimodal



Methods for assessing bimodality and trajectory shapes

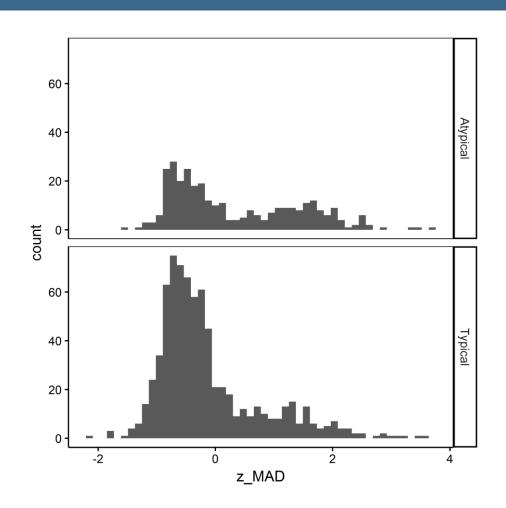
Bimodality coefficient (BC, e.g., Pfister et al., 2013)

BC =
$$\frac{m_3^2 + 1}{m_4 + 3 \cdot \frac{(n-1)^2}{(n-2)(n-3)}}$$

- Bimodal, if **BC** > **0.555**
- Hartigan's dip statistic (HDS, Hartigan & Hartigan, 1985)
 - Statistical test (H0: Distribution is unimodal)
 - If p < .05, distribution is multimodal (i.e., at least bimodal)

Assessment of bimodality

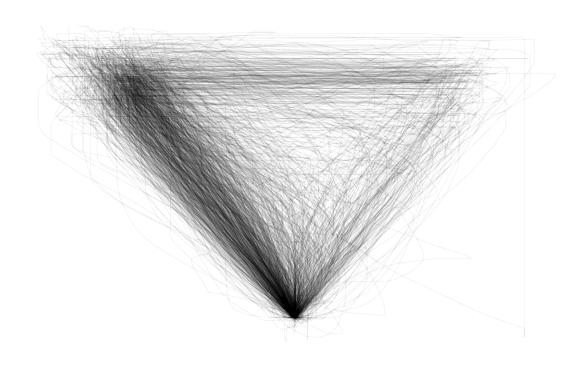
- Distribution of standardized MAD
- □ Bimodality coefficient (*BC*)
 - $BC_{\text{typical}} = .61$; $BC_{\text{atypical}} = .59$
 - Indicates bimodality as BC > .555
- Also influenced by setup of study (cf. design factors)



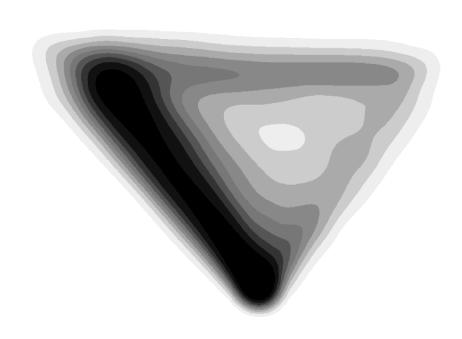
Assessing distribution of individual trajectory shapes

- Assess distribution of trajectory shapes (Wulff et al., in press)
 - Bimodality analyses so far focus on a single measure only
 - New analyses proposed taking complete **trajectory shape** into account
 - General question: is aggregate trajectory representative of individual trajectories
 - or are there **different types** of trajectories on the trial level?
- Visualization tools
 - Animations
 - Heatmaps and difference maps
- Analyses tools
 - Clustering
 - Prototype allocation

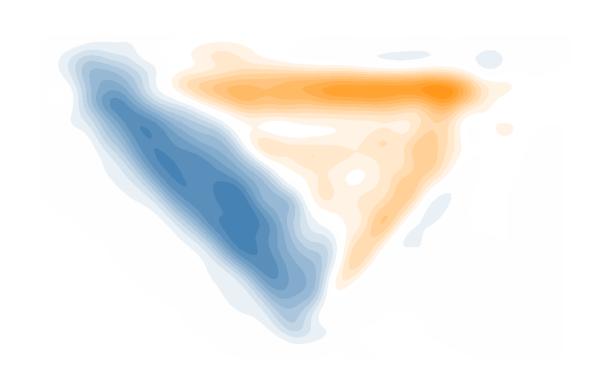
Heatmap of raw trajectories



Heatmap of raw trajectories (smoothed)



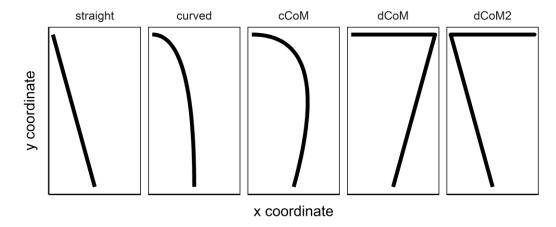
Difference map for typical vs. atypical condition



Prototype recognition (Wulff et al., in press)

Specify set of prototypes

Set of prototypes based on clustering results of the meta-analysis by Wulff et al. (in prep.)



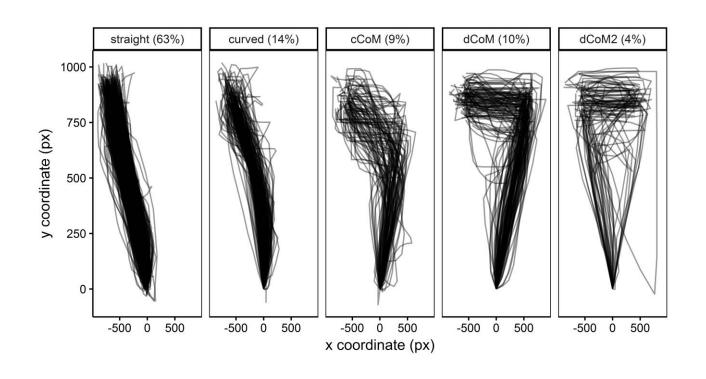
Spatialize trajectories

Resample trajectories to small number of points distributed equally across space

Assign trajectories to prototypes

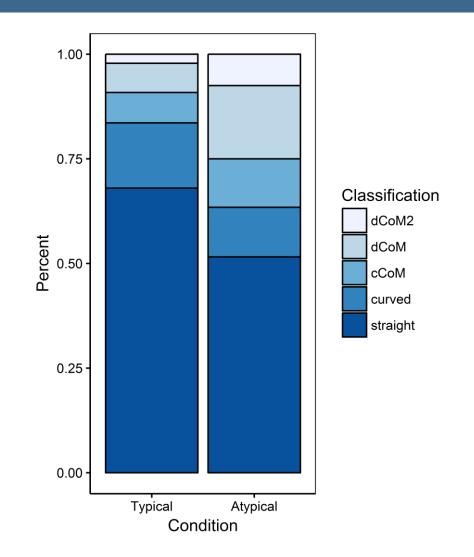
- Compute dissimilarity between every trajectory and prototype
- Assign trajectory to prototype with smallest distance
- (Potentially exclude trajectories where smallest distance is too large)

Prototype allocation for replication experiment



Classification frequencies per condition

- Relative frequency of prototype classification differs for conditions
 - $\chi^2 = 57.9, p < .001$
- Atypical condition predicts occurrence of types that indicate more conflict
 - in ordinal mixed regression model on trial level
 - with random intercept per participant
 - z = 6.74, p < .001





Previous applications

Focusing on JDM research

Mouse-tracking

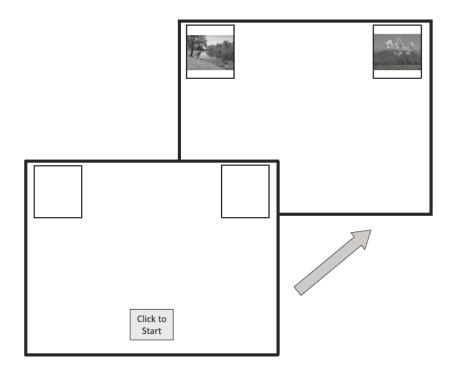
Application domains

- Application of mouse-tracking in a growing number of psychological domains (Reviews by Freeman, in press; Stillman et al., 2018)
 - Semantic processing (e.g., Spivey et al., 2005; Dale & Duran, 2011)
 - Social cognition (e.g., Freeman et al., 2008; Freeman & Ambady, 2011)
 - Learning and memory (e.g., Dale et al., 2008; Koop & Criss, 2016)
 - Self-control (e.g., Sullivan et al., 2015; Stillman et al., 2017)
- In the last years also extended to JDM research
 - Intertemporal choice (Dshemuchadse et al., 2013)
 - Moral dilemmas (Koop, 2013)
 - Decisions under risk (Koop & Johnson, 2013)
 - Social dilemmas (Kieslich & Hilbig, 2014)
 - Judgmental biases (Szaszi et al., 2018; Travers et al., 2016)

Preferential decision making

Validation experiment (Koop & Johnson, 2013, Exp. 1)

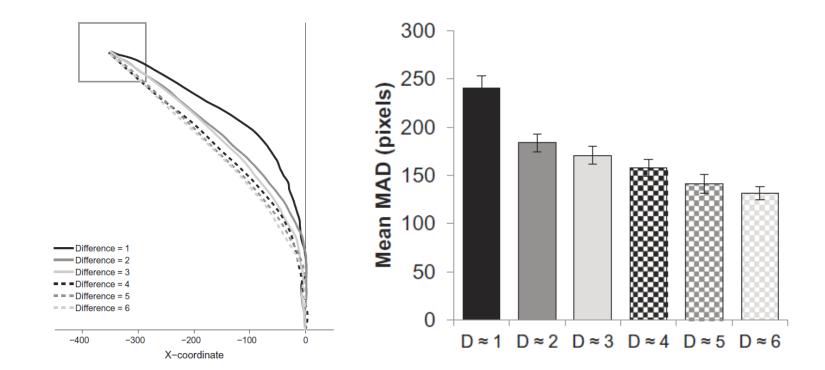
- Decisions between affective images
 - Task: Which of two images do you prefer?
 - Pictures from IAPS database: provides norms for pleasantness ratings
 - Creation of pairs where difference in preferences is systematically varied



Preferential decision making

Validation experiment (Koop & Johnson, 2013, Exp. 1)

- Decisions between affective images
 - Increase in difference of a priori preference ratings leads to
 - Decrease in trajectory curvature
 - Decrease in maximum absolute deviation (MAD)



Basic structure

- □ Social dilemma (Dawes 1980; Van Lange et al., 2013)
 - Individuals can choose between two options
 - Defection
 - Cooperation
- □ Standard social dilemma: **Prisoner's dilemma game** (PDG; Rapoport & Chammah, 1965)

		Play	er 1
		cooperates	defects
Player 2	cooperates	100 100	200 0
	defects	0 200	50 50

Spontaneous cooperation?

- □ Theoretical proposition (Rand et al., 2012, 2014)
 - People are spontaneously inclined to cooperate
 - Defection requires effortful deliberation
- Empirical test using response times
 - Idea: spontaneous = fast, deliberative = slow
 - Mixed results (e.g., Rand et al., 2014; meta-analysis by Rand, 2016; Registered replication report, 2017)
 - Other factors may influence speed (e.g., guessing, information search)
- Experiment using mouse-tracking (Kieslich & Hilbig, 2014)
 - When deciding to defect, mouse movements should be more curved towards non-chosen option (cooperation)
 - When deciding to **cooperate**, mouse movements should be less curved towards non-chosen option (defection)

Mouse-tracking experiment (Kieslich & Hilbig, 2014)

- □ Lab experiment (N = 115)
 - at the University of Mannheim
 - implementation in OpenSesame (Mathôt et al., 2012) in combination with
 - mousetrap plug-ins for mouse-tracking (Kieslich & Henninger, 2017)
 - psynteract plug-ins for interactive experiments (Henninger, Kieslich, & Hilbig, 2017)
- Participants play 15 two-person social dilemma games
 - without receiving feedback
 - random order
 - incentivized (5 interactions paid out, Ø payout: 2.56 €)
- Social dilemma games
 - 5 x prisoner's dilemma game (PDG)
 - 5 x chicken game
 - 5 x stag hunt game

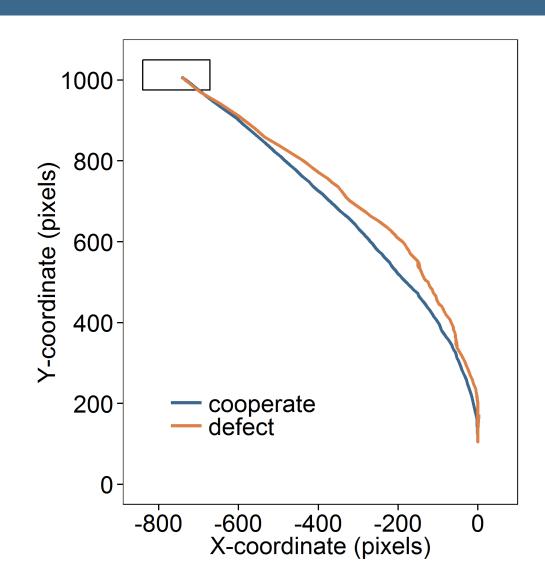
Decision 9 of 15

	You choose Option A	You choose Option B	
Person 2 chooses Option A	100 100	200 0	Person 2 chooses Option A
Person 2 chooses Option B	0 200	50 50	Person 2 chooses Option B
	You choose Option A	You choose Option B	

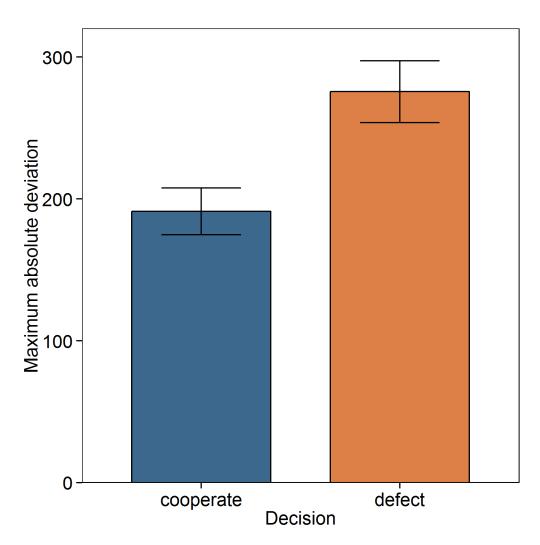
Please choose between Option A and B.

Start

Average time-normalized response trajectories



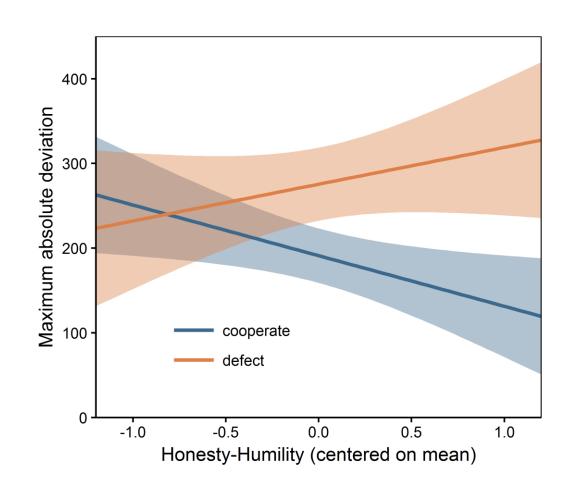
Maximum deviation per decision



- Main effect of decision
 - MAD significantly higher for defection than for cooperation
- Effect replicated
 - With different measures
 - With filtered trials
 - With linear mixed model on trial level

Predicting individual differences in conflict

- Individual differences in conflict: Differences should be stronger for individuals high in Honesty-Humility
 - Dispositional cooperativeness
 - Basic personality factor in the HEXACO personality model (Ashton & Lee, 2007)
- Significant interaction between HH and decision



Mouse-tracking challenges

Experimental control over comparison dimension

- Mouse-tracking tasks usually involves "correct"/desired response option
 + comparison dimension is experimentally manipulated
- Here final choice constitutes comparison dimension of interest
 - loss of experimental control
 - use of different games to achieve variation in cooperation rates

Complexity and amount of information

- Amount of information and complexity of decision considerably higher than in previous tasks
- Mouse movements more noisy (e.g., reading movements in some trials)
 - Current solution: analyses replicated with and without problematic trials
 - Ideal solution: simpler task design with less information
 → working on conceptual replication in binary public goods game, also taking into account the newly proposed analytical approaches (prototype mapping)

Action selection

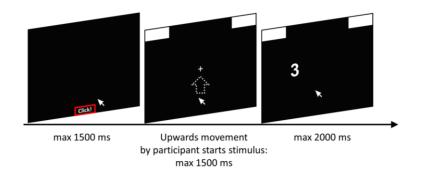
Simon effect and conflict adaptation (Scherbaum et al., 2010)

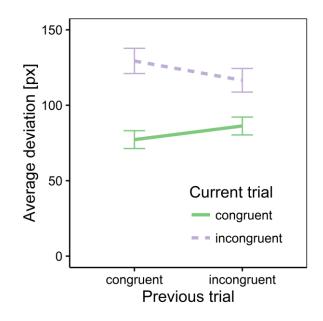
Mouse-tracking in Simon task

- Participants have to click on left vs. right option depending on the stimulus
 (e.g., left if number < 5, otherwise right)
- Position of stimulus varied (left vs. right) so that desired response and position are either congruent or incongruent

Results

- Simon effect: larger deviations in incongruent than in congruent trials
- Conflict adaptation: Simon effect reduced if previous trial was incongruent





Reanalysis from Scherbaum & Kieslich (in press)

Action selection

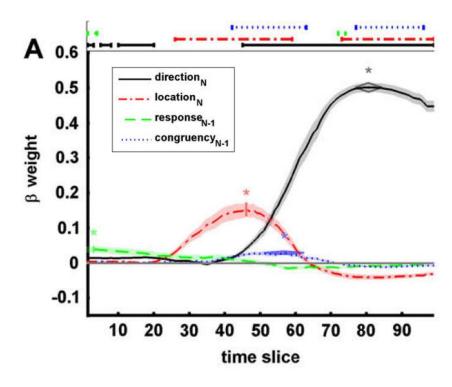
Time continuous multiple regression (Scherbaum et al., 2010)

Time continuous multiple regression

- □ Criterion: mouse movement angles on the XY plane (≈ movement direction)
- Separate regressions per time step and participant
- Reveals temporal order and strength with which each predictor influences preference development

Predictors

- Task relevant
 - Direction (left/right)
- Task irrelevant
 - stimulus location (left/right)
 - previous response (left/right)
 - congruency sequence (same/different)



Average β weights per time step and predictor.

Design factors

Overview

- Researchers face a number of design choices
 when creating mouse-tracking experiments
 - Starting procedure (static, restricted initiation time, dynamic)
 - Cursor speed settings (velocity & acceleration)
 - Indicate response via click vs. touch
- Some authors have given recommendations about designing mouse-tracking studies (Fischer & Hartmann, 2014; Hehman et al., 2015)
- Empirical validation studies are being conducted (Scherbaum & Kieslich, in press; Kieslich et al., in preparation)

Design factors

Preliminary summary of findings

Response indication

□ Click on button leads to larger effects than touch – effect related to higher proportion of trials with extreme movements to non-chosen option

Mouse sensitivity settings

- Did not significantly influence effect of interest in static setup although default settings generally lead to more extreme curvature than reduced mouse speed
- Reducing mouse speed becomes relevant for dynamic start condition to ensure stimulus information can be acquired during upwards movement

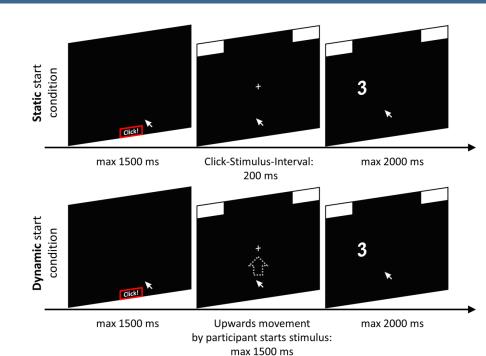
Starting procedure

- Restricting maximum initiation time led to larger effects a dynamic start or restricting maximum response time only influenced shape but not effect size
- However, restricting initiation times also led to largest proportion of excluded trials (and seemed to be challenging for some participants)

Method (Scherbaum & Kieslich, in press)

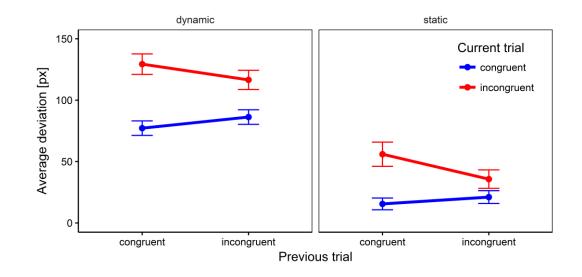
Mouse-tracking in Simon task

- Participants click on left vs. right option depending on stimulus (left if number < 5, otherwise right)
- Position of stimulus varied (left vs. right) so that desired response and position are either congruent or incongruent
- Variation starting procedure
 - Dynamic: move upwards to display stimulus (data from Scherbaum et al., 2010)
 - Static: stimulus displayed after fixed interval of 200 ms (typical duration of movement initiation in dynamic condition) (new data)



Discrete effects: Results for average deviation

- Simon effect and congruency sequence effect replicated in both conditions
- No significant interaction of theoretically important effects with starting procedure



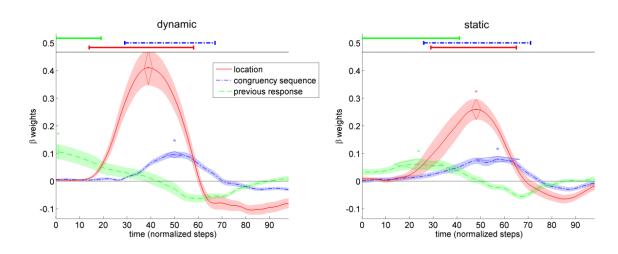
Error bars represent 1 SEM.

Dynamic effects: Time-continuous angle regression

 Time continuous multiple regression predicting vertical movement angle at each time point

Predictors

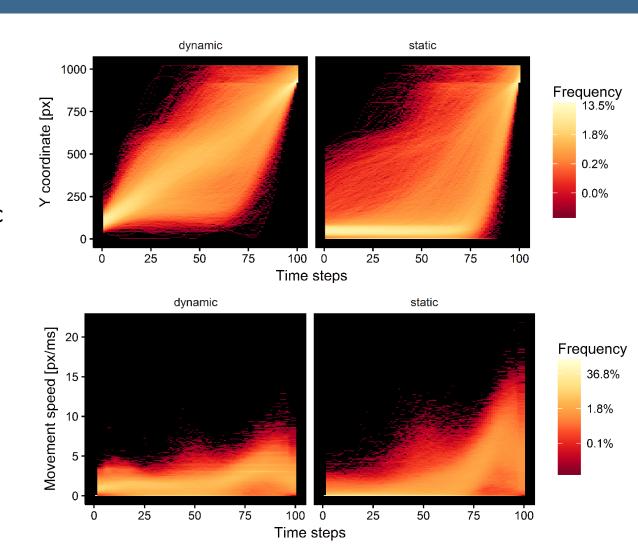
- location (congruency)
- congruency sequence (same / different)
- previous response (same / different)
- Effects stronger and more temporarily distinct in dynamic starting condition



Average β weights per time step and predictor. Lines indicate segments of β weights significantly > 0.

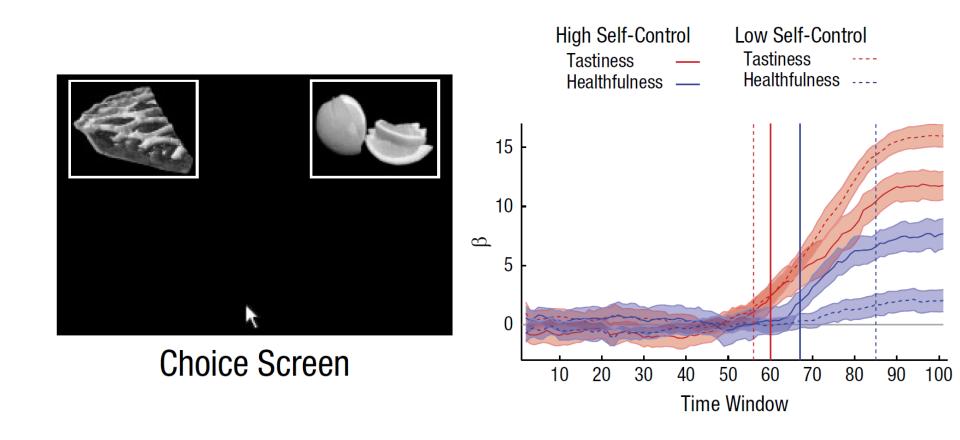
Movement consistency

- Smooth and consistent upwards movement in dynamic starting condition
- Participants in static starting condition often stay at bottom of screen for more than half of the trial before moving upwards quickly



Self-control

Food choices (Sullivan et al., 2015)



Decisions under risk

Basic paradigm

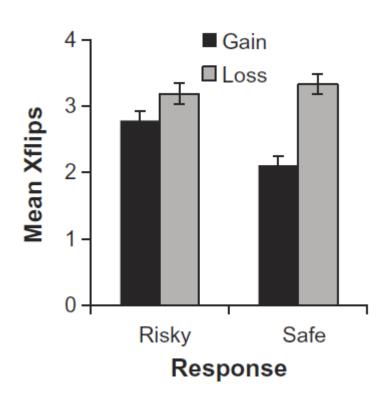
- Risky choice / decisions under risk
 - Which of the two gambles do you want to play?

Gamble A	Gamble B
You have a 50% chance of winning \$90, otherwise nothing	You have a 90% chance of winning \$50, otherwise nothing

- □ Gamble A: "risky"
 - Higher amount, lower probability of winning
- Gamble B: "safe"
 - Lower amount, higher probability of winning

Decisions under risk

x-flips (Koop & Johnson, 2013, Exp. 2)

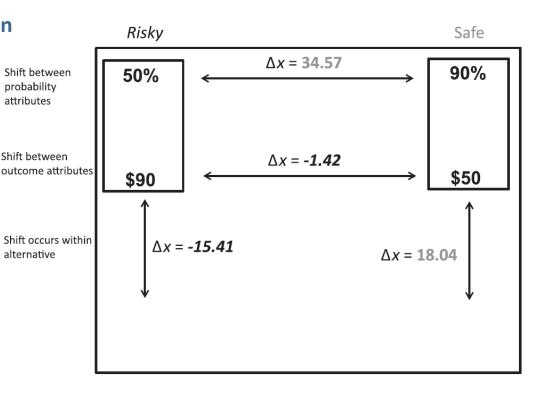


Decisions under risk

Combining mouse- and eye-tracking (Koop & Johnson, 2013, Exp. 3)

attributes

- Change in x-position (Δx) as function of transitions of attention
 - $\triangle x > 0$: movement towards safe gamble
 - $\Delta x < 0$: movement towards risky gamble
- Evidence accumulation model
 - Predict momentary preference based on visual input
 - Mean correlation between predicted preference and x-position is r = .78
- **Conclusions**
 - Visual attention to probability and outcome information predicts mouse response
 - Mouse movements largely reflect quality of acquired information





Your experiments

Mouse-tracking introduction (Monday)

- 13:00-14:30 General introduction to mouse-tracking
 - Paradigm and assumptions
 - Implementation and analysis
 - Previous applications
- 14:30-15:00 Introduction to task
 - Type of experiments considered
 - Your tasks during the workshop
- 15:00-17:00 Develop experimental design conceptually
- □ 17:00-18:00 Present experimental design in plenum

Your tasks during the workshop

- Goal of workshop
 - Design, build, pre-register, run, and analyze a mouse-tracking experiment
 - In small groups
- Monday
 - Develop experimental design (task, manipulation, hypotheses, measures)
 - Present experimental design in plenum
- Tuesday
 - Build experiment
 - Register experiment at OSF
 - Participate in experiments
- Wednesday
 - Analyze and visualize your data
 - Discuss your results
- Saturday
 - Present results

Type of experiments

- In the experiment, participants complete a number of trials that involve decisions of the same structure
- In each trial, participants have to decide between two options by clicking on the corresponding button (two-alternative forced choice task, 2AFC)
- Between trials, the stimulus to be decided upon varies (usually)
 and / or the two response categories
- The stimulus (and/or the response options in case they vary) should be simple (e.g., a single word, a picture)

Implementation & analysis

Software

- Custom extensions for experimental software
 - Code based implementations, e.g., in E-Prime or MATLAB
 - Also need scripts for preprocessing the data
 - Require programming skills
- MouseTracker (Freeman & Ambady, 2010)
 - Stand-alone program
 - Relatively easy to use, but limited in features and flexibility
 - Free of charge but closed source, Windows only



- □ Mousetrap (Kieslich & Henninger, 2017; Kieslich, Wulff et al., in preparation)
 - Drag & drop plugins for experimental software OpenSesame
 - R package mousetrap for preprocessing and analysis
 - Open source, free of charge, cross-platform
 - Available from http://pascalkieslich.github.io/mousetrap/





Software for the workshop

- To create mouse-tracking experiments, first install OpenSesame. It is available from http://osdoc.cogsci.nl/3.2/download/.
- To install the mousetrap plugin for OpenSesame, follow the instructions at https://github.com/pascalkieslich/mousetrap-os#installation. Please make sure to install the latest version of OpenSesame (3.2.4) and the development version of the mousetrap-os plugin.
- To analyze mouse-tracking data install R (https://www.r-project.org/) and RStudio (https://www.rstudio.com/products/rstudio/download/).
- Afterwards, please run the following command in R to install the required packages: install.packages(c("readbulk", "mousetrap"))

Thank you!

Questions and comments are highly appreciated!

Now & via email: <u>kieslich@psychologie.uni-mannheim.de</u> <u>dirk.wulff@gmail.com</u>

Mousetrap-os plugins: https://github.com/pascalkieslich/mousetrap-os

Mousetrap R package: http://pascalkieslich.github.io/mousetrap/

Thanks:

Felix Henninger, co-developer of mousetrap-os plugin and R package Jonas Haslbeck & Michael Schulte-Mecklenbeck, co-developers of mousetrap R package

Mila Rüdiger and Monika Wiegelmann for data collection and testing